U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____

Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 11-04-02
Searcher: Beverly e 4994
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 25
Number of Searches: _____
Number of Databases: 1

Search Site

_____ STIC
_____ CM-1
_____ Pre-S

Type of Search

_____ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

Vendors

_____ IG Suite
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
✓ Other CGN

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:08:53 ; Search time 42.6923 Seconds
(without alignments)
10293.594 Million cell updates/sec

Title: US-09-981-606-16
Perfect score: 21
Sequence: 1 acaagacctcagacttccagc 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	21	100.0	249	9	HSU80914	U80914 Human hered
c 2	21	100.0	874	9	HSHLAH2	Y09800 H.sapiens H
c 3	21	100.0	10825	6	AR117789	AR117789 Sequence
c 4	21	100.0	10825	6	AR117790	AR117790 Sequence
c 5	21	100.0	10825	6	AR117791	AR117791 Sequence
c 6	21	100.0	10825	6	AR117792	AR117792 Sequence
c 7	21	100.0	10825	6	AR149459	AR149459 Sequence
c 8	21	100.0	10825	6	AR149460	AR149460 Sequence
c 9	21	100.0	10825	6	AR149461	AR149461 Sequence
c 10	21	100.0	10825	6	AR149462	AR149462 Sequence
c 11	21	100.0	12146	9	HSHFE	Z92910 Homo sapien
12	21	100.0	193752	2	AL359892	AL359892 Homo sapi
c 13	21	100.0	246240	6	AR036572	AR036572 Sequence
c 14	21	100.0	246240	6	AR036573	AR036573 Sequence
c 15	21	100.0	246240	6	AR036574	AR036574 Sequence
16	21	100.0	246282	9	HSU91328	U91328 Human hered
c 17	19.4	92.4	177562	9	HS269M15	AL021395 Human DNA
18	18.4	87.6	1552	8	AF254643	AF254643 Penicilli
19	18.4	87.6	1649	8	AF284645	AF284645 Aspergill
20	18.4	87.6	164920	9	AC093908	AC093908 Homo sapi
c 21	18.4	87.6	166651	4	AC087160	AC087160 Sus scrof
c 22	18.4	87.6	198614	10	AL596127	AL596127 Mouse DNA
c 23	18.4	87.6	209111	9	CNS00YVG	AL096821 Human chr
24	17.8	84.8	879	4	AF268466	AF268466 Bos tauru
c 25	17.8	84.8	1975	10	MMU69491	U69491 Mus musculu
c 26	17.8	84.8	1985	10	MMIL11	X98519 M.musculus
c 27	17.8	84.8	68582	9	AB017652	AB017652 Homo sapi
28	17.8	84.8	68582	9	AP000464	AP000464 Homo sapi
c 29	17.8	84.8	102209	2	AC106935	AC106935 Rattus no
c 30	17.8	84.8	110665	9	HS696P19	AL035588 Human DNA
c 31	17.8	84.8	153101	2	AC026818	AC026818 Homo sapi
32	17.8	84.8	167901	9	AP002991	AP002991 Homo sapi
c 33	17.8	84.8	168473	2	AC096172	AC096172 Rattus no
c 34	17.8	84.8	173952	2	AC021903	AC021903 Homo sapi
c 35	17.8	84.8	178780	9	AC069302	AC069302 Homo sapi
c 36	17.8	84.8	179151	9	AC024909	AC024909 Homo sapi
37	17.8	84.8	179642	2	AC099288	AC099288 Rattus no
38	17.8	84.8	180442	9	AC009549	AC009549 Homo sapi
39	17.8	84.8	196591	2	AL589668	AL589668 Homo sapi
c 40	17.8	84.8	208543	2	AC100736	AC100736 Mus muscu
41	17.8	84.8	209350	2	AC106531	AC106531 Rattus no
c 42	17.8	84.8	256501	2	AC090655	AC090655 Mus muscu
43	17.4	82.9	614	10	AB002393	AB002393 Rattus no
c 44	17.4	82.9	58020	2	AP002772	AP002772 Homo sapi
45	17.4	82.9	153149	9	AL354819	AL354819 Human DNA

ALIGNMENTS

RESULT 1
HSU80914/c
LOCUS HSU80914 249 bp DNA linear PRI 05-JAN-1999
DEFINITION Human hereditary haemochromatosis protein (HLA-H) gene, partial cds.
ACCESSION U80914
VERSION U80914.1 GI:4098856
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 249)
AUTHORS Hashimoto,K., Hirai,M. and Kurosawa,Y.
TITLE Identification of a mouse homolog for the human hereditary haemochromatosis candidate gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 249)
AUTHORS Hashimoto,K.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-1996) Institute for Comprehensive Medical Science, Fujita Health University, Aichi, Toyoake 470-11, Japan
FEATURES Location/Qualifiers

```
     source          1. .249
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22"
                     /tissue_type="placenta"
     mRNA            join(<1. .20,230. .>249)
                     /gene="HLA-H"
     gene            <1. .>249
                     /gene="HLA-H"
     CDS             join(<1. .20,230. .>249)
                     /gene="HLA-H"
                     /note="MR2(alpha 1/intron/alpha 2)"
                     /codon_start=2
                     /product="hereditary haemochromatosis protein"
                     /protein_id="AAD00449.1"
                     /db_xref="GI:4098857"
                     /translation="NHNHSKESHTLQV"
BASE COUNT       59 a     50 c     75 g     65 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 9;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db    115 ACAAGACCTCAGACTTCCAGC 95


RESULT 2
HSHLAH2/c
LOCUS       HSHLAH2                  874 bp    DNA     linear   PRI 23-JUL-1999
DEFINITION  H.sapiens HFE gene, exon 2 & 3.
ACCESSION   Y09800
VERSION     Y09800.1  GI:2370112
KEYWORDS    HFE gene.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 874)
  AUTHORS   Carella and Gasparini,P.
  TITLE     Hereditary hemochromatosis genomic structure and organization of
            HLA-H gene
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 874)
  AUTHORS   Gasparini,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1996) P. Gasparini, Servizio de Genetica Medica -
            IRCCS, 'Ospedale CSS', Via Cappuccini, 71013 S Giovanni, Rotondo
            (FG), ITALY
COMMENT     Related sequence: U60319.
FEATURES             Location/Qualifiers
     source          1. .874
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="31H6"
                     /map="6p22"
     exon            46. .309
                     /gene="HFE"
                     /label=ex2
                     /usedin=Y09801:hfe_cds
                     /usedin=Y09801:hfe_mrna
     gene            46. .793
                     /gene="HFE"
     exon            518. .793
                     /gene="HFE"
                     /label=ex3
                     /usedin=Y09801:hfe_cds
                     /usedin=Y09801:hfe_mrna
BASE COUNT      202 a    194 c    267 g    211 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 9;  Length 874;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db    403 ACAAGACCTCAGACTTCCAGC 383


RESULT 3
AR117789/c
LOCUS       AR117789               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6140305.
ACCESSION   AR117789
VERSION     AR117789.1  GI:14098695
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 1 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a   2253 c   2648 g   2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db   4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 4
AR117790/c
LOCUS       AR117790               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 3 from patent US 6140305.
ACCESSION   AR117790
VERSION     AR117790.1  GI:14098696
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 3 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a   2253 c   2647 g   2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db   4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 5
AR117791/c
LOCUS       AR117791               10825 bp    DNA     linear   PAT 16-MAY-2001
```

```
DEFINITION  Sequence 5 from patent US 6140305.
ACCESSION   AR117791
VERSION     AR117791.1  GI:14098697
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 5 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a   2252 c   2649 g   2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db   4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 6
AR117792/c
LOCUS       AR117792               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 7 from patent US 6140305.
ACCESSION   AR117792
VERSION     AR117792.1  GI:14098698
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 7 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a   2252 c   2648 g   2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db   4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 7
AR149459/c
LOCUS       AR149459               10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 1 from patent US 6228594.
ACCESSION   AR149459
VERSION     AR149459.1  GI:15114050
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 1 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a   2253 c   2648 g   2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db   4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 8
AR149460/c
LOCUS       AR149460               10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 3 from patent US 6228594.
ACCESSION   AR149460
VERSION     AR149460.1  GI:15114051
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 3 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a   2253 c   2647 g   2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db   4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 9
AR149461/c
LOCUS       AR149461               10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 5 from patent US 6228594.
ACCESSION   AR149461
VERSION     AR149461.1  GI:15114052
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 5 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a   2252 c   2649 g   2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
```

```
          ||||||||||||||||||||
Db    4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 10
AR149462/c
LOCUS       AR149462               10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 7 from patent US 6228594.
ACCESSION   AR149462
VERSION     AR149462.1  GI:15114053
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 7 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a   2252 c   2648 g   2926 t
ORIGIN

  Query Match           100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity  100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
          ||||||||||||||||||||
Db    4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 11
HSHFE/c
LOCUS       HSHFE                  12146 bp    DNA     linear   PRI 23-JUL-1999
DEFINITION  Homo sapiens HFE gene.
ACCESSION   Z92910
VERSION     Z92910.1  GI:1890179
KEYWORDS    haemochromatosis; HFE gene.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 858)
  AUTHORS   Albig,W., Drabent,B., Burmester,N., Bode,C. and Doenecke,D.
  TITLE     The haemochromatosis candidate gene HFE (HLA-H) of man and mouse is
            located in syntenic regions within the histone gene cluster
  JOURNAL   J. Cell. Biochem. 69 (2), 117-126 (1998)
  MEDLINE   98208340
REFERENCE   2  (bases 1 to 12146)
  AUTHORS   Albig,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-1997) Albig W., Georg-August-Universitaet
            Goettingen, Biochemie und Molekulare Zellbiologie, Humboldtallee
            23, Goettingen, FRG, 37073
FEATURES             Location/Qualifiers
     source          1. .12146
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p"
                     /clone="ICRFy901D1223"
                     /clone_lib="ICRF YAC-library"
     exon            1028. .1324
                     /gene="HFE"
                     /number=1
     gene            1028. .10637
                     /gene="HFE"
     CDS             join(1249. .1324,4652. .4915,5125. .5400,6494. .6769,
                     6928. .7041,7995. .8035)
                     /gene="HFE"
                     /function="iron metabolism"
                     /note="haemochromatosis candidate gene"
                     /codon_start=1
                     /protein_id="CAB07442.1"
                     /db_xref="GI:1890180"
                     /db_xref="SWISS-PROT:Q30201"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                     SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
                     WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                     RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTS
                     SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGE
                     EQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQG
                     SRGAMGHYVLAERE"
     intron          1325. .4651
                     /gene="HFE"
                     /number=1
     repeat_unit     3494. .3735
                     /gene="HFE"
                     /rpt_family="Alu"
     repeat_unit     3973. .4283
                     /gene="HFE"
                     /rpt_family="Alu"
     exon            4652. .4915
                     /gene="HFE"
                     /number=2
     intron          4916. .5124
                     /gene="HFE"
                     /number=2
     exon            5125. .5400
                     /gene="HFE"
                     /number=3
     intron          5401. .6493
                     /gene="HFE"
                     /number=3
     repeat_unit     5707. .6005
                     /gene="HFE"
                     /rpt_family="Alu"
     exon            6494. .6769
                     /gene="HFE"
                     /number=4
     intron          6770. .6927
                     /gene="HFE"
                     /number=4
     exon            6928. .7041
                     /gene="HFE"
                     /number=5
     intron          7042. .7994
                     /gene="HFE"
                     /number=5
     exon            7995. .9050
                     /gene="HFE"
                     /number=6
     repeat_unit     9017. .9340
                     /gene="HFE"
                     /rpt_family="Alu"
     intron          9051. .10205
                     /gene="HFE"
                     /number=6
     repeat_unit     9957. .10239
                     /gene="HFE"
                     /rpt_family="Alu"
     exon            10206. .10637
                     /gene="HFE"
                     /number=7
     polyA_signal    10617. .10622
                     /gene="HFE"
BASE COUNT     3383 a   2474 c   2911 g   3378 t
ORIGIN

  Query Match           100.0%;  Score 21;  DB 9;  Length 12146;
  Best Local Similarity  100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db   5010 ACAAGACCTCAGACTTCCAGC 4990


RESULT 12
AL359892
LOCUS       AL359892              193752 bp    DNA     linear   HTG 13-JUN-2001
DEFINITION  Homo sapiens chromosome 6 clone RP11-557F22, *** SEQUENCING IN
            PROGRESS ***, 18 unordered pieces.
ACCESSION   AL359892
VERSION     AL359892.5  GI:9930971
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 193752)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 27, 2000 this sequence version replaced gi:9864230.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA557F22
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 183925 bases at least Q40
            Consensus quality: 187703 bases at least Q30
            Consensus quality: 189658 bases at least Q20
            Insert size: 192052; sum-of-contigs
            Insert size: 198247; agarose-fp
            Quality coverage: 3.68x in Q20 bases; sum-of-contigs Quality
            coverage: 3.70x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 18 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3250: contig of 3250 bp in length
            *     3251 3350: gap of      100 bp
            *     3351    14600: contig of 11250 bp in length
            *    14601 14700: gap of      100 bp
            *    14701    32357: contig of 17657 bp in length
            *    32358 32457: gap of      100 bp
            *    32458    34886: contig of 2429 bp in length
            *    34887 34986: gap of      100 bp
            *    34987    43490: contig of 8504 bp in length
            *    43491 43590: gap of      100 bp
            *    43591    47437: contig of 3847 bp in length
            *    47438 47537: gap of      100 bp
            *    47538    57356: contig of 9819 bp in length
            *    57357 57456: gap of      100 bp
            *    57457    59845: contig of 2389 bp in length
            *    59846 59945: gap of      100 bp
            *    59946    63972: contig of 4027 bp in length
            *    63973 64072: gap of      100 bp
            *    64073    82711: contig of 18639 bp in length
            *    82712 82811: gap of      100 bp
            *    82812   111814: contig of 29003 bp in length
            *   111815 111914: gap of      100 bp
            *   111915   120276: contig of 8362 bp in length
            *   120277 120376: gap of      100 bp
            *   120377   136660: contig of 16284 bp in length
            *   136661 136760: gap of      100 bp
            *   136761   153913: contig of 17153 bp in length
            *   153914 154013: gap of      100 bp
            *   154014   158659: contig of 4646 bp in length
            *   158660 158759: gap of      100 bp
            *   158760   164235: contig of 5476 bp in length
            *   164236 164335: gap of      100 bp
            *   164336   184996: contig of 20661 bp in length
            *   184997 185096: gap of      100 bp
            *   185097   193752: contig of 8656 bp in length.
FEATURES             Location/Qualifiers
     source          1. .193752
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-557F22"
                     /clone_lib="RPCI-11.2"
     misc_feature    1. .3250
                     /note="assembly_fragment:01752
                     fragment_chain:1"
     misc_feature    3351. .14600
                     /note="assembly_fragment:01177
                     fragment_chain:1"
     misc_feature    14701. .32357
                     /note="assembly_fragment:00673
                     fragment_chain:1"
     misc_feature    32458. .34886
                     /note="assembly_fragment:00884
                     fragment_chain:1"
     misc_feature    34987. .43490
                     /note="assembly_fragment:00652
                     fragment_chain:1"
     misc_feature    43591. .47437
                     /note="assembly_fragment:00628
                     fragment_chain:1"
     misc_feature    47538. .57356
                     /note="assembly_fragment:00983
                     fragment_chain:2"
     misc_feature    57457. .59845
                     /note="assembly_fragment:01129
                     fragment_chain:2"
     misc_feature    59946. .63972
                     /note="assembly_fragment:01207
                     fragment_chain:3"
     misc_feature    64073. .82711
                     /note="assembly_fragment:00301
                     fragment_chain:3"
     misc_feature    82812. .111814
                     /note="assembly_fragment:00470"
     misc_feature    111915. .120276
                     /note="assembly_fragment:00783.0"
     misc_feature    120377. .136660
                     /note="assembly_fragment:01014"
     misc_feature    136761. .153913
                     /note="assembly_fragment:01161"
     misc_feature    154014. .158659
                     /note="assembly_fragment:01329"
     misc_feature    158760. .164235
                     /note="assembly_fragment:01675"
     misc_feature    164336. .184996
                     /note="assembly_fragment:01884"
     misc_feature    185097. .193752
                     /note="assembly_fragment:01893"
BASE COUNT    55929 a  40409 c  39996 g  55686 t   1732 others
ORIGIN

  Query Match           100.0%;  Score 21;  DB 2;  Length 193752;
  Best Local Similarity  100.0%;  Pred. No. 0.99;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db  12532 ACAAGACCTCAGACTTCCAGC 12552


RESULT 13
AR036572/c
LOCUS       AR036572              246240 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 20 from patent US 5872237.
ACCESSION   AR036572
VERSION     AR036572.1  GI:5953240
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246240)
  AUTHORS   Feder,J.Nathan., Kronmal,G.Scott., Lauer,P.M., Ruddy,D.A.,
            Thomas,W., Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Megabase transcript map: novel sequences and antibodies thereto
  JOURNAL   Patent: US 5872237-A 20 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .246240
                     /organism="unknown"
BASE COUNT    73211 a  50177 c  50599 g  72252 t      1 others
ORIGIN

  Query Match            100.0%;  Score 21;  DB 6;  Length 246240;
  Best Local Similarity  100.0%;  Pred. No. 0.98;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db 196423 ACAAGACCTCAGACTTCCAGC 196403


RESULT 14
AR036573/c
LOCUS       AR036573              246240 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 21 from patent US 5872237.
ACCESSION   AR036573
VERSION     AR036573.1  GI:5953241
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246240)
  AUTHORS   Feder,J.Nathan., Kronmal,G.Scott., Lauer,P.M., Ruddy,D.A.,
            Thomas,W., Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Megabase transcript map: novel sequences and antibodies thereto
  JOURNAL   Patent: US 5872237-A 21 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .246240
                     /organism="unknown"
BASE COUNT    73211 a  50177 c  50599 g  72252 t      1 others
ORIGIN

  Query Match            100.0%;  Score 21;  DB 6;  Length 246240;
  Best Local Similarity  100.0%;  Pred. No. 0.98;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db 196423 ACAAGACCTCAGACTTCCAGC 196403


RESULT 15
AR036574/c
LOCUS       AR036574              246240 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 22 from patent US 5872237.
ACCESSION   AR036574
VERSION     AR036574.1  GI:5953242
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246240)
  AUTHORS   Feder,J.Nathan., Kronmal,G.Scott., Lauer,P.M., Ruddy,D.A.,
            Thomas,W., Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Megabase transcript map: novel sequences and antibodies thereto
  JOURNAL   Patent: US 5872237-A 22 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .246240
                     /organism="unknown"
BASE COUNT    73211 a  50177 c  50599 g  72252 t      1 others
ORIGIN

  Query Match            100.0%;  Score 21;  DB 6;  Length 246240;
  Best Local Similarity  100.0%;  Pred. No. 0.98;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db 196423 ACAAGACCTCAGACTTCCAGC 196403


Search completed: November  2, 2002, 05:41:13
Job time : 124.692 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:05:23 ; Search time 4.27747 Seconds
(without alignments)
8429.091 Million cell updates/sec

Title: US-09-981-606-16
Perfect score: 21
Sequence: 1 acaagacctcagacttccagc 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	21	100.0	21	21	AAA96783	PCR primer for his
c	2	21	100.0	5749	22	AAL36747	Human musculoskele
c	3	21	100.0	10825	18	AAT96690	Hereditary haemoch
c	4	21	100.0	10825	22	AAC68425	Human hereditary h
c	5	21	100.0	10825	22	AAC68426	Human hereditary h
c	6	21	100.0	10825	22	AAC68427	Human hereditary h
c	7	21	100.0	10825	22	AAC68428	Human hereditary h
c	8	21	100.0	12146	21	AAA96794	Genomic DNA of a h
	9	21	100.0	235033	19	AAV57926	Hereditary haemoch
	10	21	100.0	237326	19	AAV57903	Hereditary haemoch
	11	18.4	87.6	624	21	AAF11290	Aspergillus niger
	12	17.8	84.8	440	22	AAK54198	Murine transport a
c	13	17.8	84.8	485	22	ABA43533	Human breast cell
c	14	17.8	84.8	485	22	ABA53984	Human foetal liver
c	15	17.8	84.8	485	22	ABA23734	Probe #2200 for ge
c	16	17.8	84.8	485	22	AAK02253	Human brain expres
c	17	17.8	84.8	485	22	AAK27702	Human bone marrow
c	18	17.8	84.8	485	22	AAI12279	Probe #2212 for ge
c	19	17.8	84.8	485	22	AAI33635	Probe #2321 used t
c	20	17.8	84.8	485	22	AAI02195	Probe #2186 used t
	21	16.8	80.0	969	22	ABA20561	Human nervous syst
	22	16.8	80.0	969	22	ABA20562	Human nervous syst
	23	16.8	80.0	2816	23	AAS93899	DNA encoding novel
	24	16.8	80.0	10091	22	AAK69350	Human immune/haema
	25	16.8	80.0	30032	22	ABA17086	Human nervous syst
	26	16.8	80.0	52562	22	AAK86669	Human immune/haema
	27	16.8	80.0	53075	22	AAK86671	Human immune/haema
	28	16.4	78.1	1482	23	AAS67663	DNA encoding novel
c	29	16.4	78.1	3249	16	AAQ82748	pRb2 retinoblastom
c	30	16.4	78.1	3291	23	AAS83193	DNA encoding novel
c	31	16.4	78.1	6795	23	AAS83194	DNA encoding novel
	32	16.2	77.1	684	22	AAH03778	Human cDNA clone (
c	33	16.2	77.1	700	22	AAH92923	Human inflammatory
c	34	16.2	77.1	700	22	AAH92924	Human inflammatory
c	35	16.2	77.1	1215	18	AAT76885	Arabidopsis floral
c	36	16.2	77.1	1215	19	AAV58306	Arabidopsis thalia
c	37	16.2	77.1	1215	21	AAC61407	cDNA encoding a AP
c	38	16.2	77.1	1705	17	AAT17868	Murine interleukin
c	39	16.2	77.1	1714	17	AAT32613	Murine Etl-2 gene.
c	40	16.2	77.1	1714	22	AAS11971	Mouse cDNA encodin
	41	16.2	77.1	2786	21	AAC64792	Human plakophilin-
	42	16.2	77.1	2818	22	AAK52359	Human polynucleoti
c	43	16.2	77.1	2832	23	AAS72205	DNA encoding novel
c	44	16.2	77.1	2837	22	AAK53343	Human polynucleoti
c	45	16.2	77.1	3386	23	AAS86846	DNA encoding novel

ALIGNMENTS

RESULT 1
AAA96783
ID AAA96783 standard; DNA; 21 BP.
XX
AC AAA96783;
XX
DT 19-FEB-2001 (first entry)
XX
DE PCR primer for histocompatibility iron loading (HFE) gene exon 2.
XX
KW Human; histocompatibility iron loading protein; HFE protein;
KW major histocompatibility complex; non-classical class I gene;
KW chromosome 6p; iron disorder; haemochromatosis; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200058515-A1.
XX
PD 05-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US07982.
XX
PR 26-MAR-1999; 99US-0277457.
XX
PA (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI Rothenberg BE, Sawada-Hirai R, Barton JC;
XX
DR WPI; 2000-647244/62.
XX
PT Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT susceptibility to develop it, by determining the presence of a mutation

PT in exon 2 or an intron of a histocompatibility iron loading nucleic
PT acid -
XX
PS Claim 24; Page 5; 55pp; English.
XX
CC PCR primers A96782-83 were used to amplify a fragment of the human
CC histocompatibility iron loading (HFE) gene. The HFE gene is a major
CC histocompatibility (MHC) non-classical class I gene located on
CC chromosome 6p. Mutations in the gene lead to iron disorders. The
CC specification describes a method for diagnosing an iron disorder or a
CC genetic susceptibility to develop the disorder in a mammal. The method
CC comprises determining the presence of a mutation in exon 2 or an intron
CC of a HFE gene or protein. The mutation is not a C to G missense mutation
CC at nucleotide 187 of the sequence given in A96769 (Genbank Accession
CC number U60319). The presence of the mutation indicates the disorder or
CC the genetic susceptibility to the disorder. The method is used to
CC diagnose an iron disorder e.g. haemochromatosis, or a genetic
CC susceptibility to develop it.
XX
SQ Sequence 21 BP; 7 A; 8 C; 3 G; 3 T; 0 other;

Query Match 100.0%; Score 21; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
||||||||||||||||||||
Db 1 ACAAGACCTCAGACTTCCAGC 21

RESULT 2
AAL36747/c
ID AAL36747 standard; DNA; 5749 BP.
XX
AC AAL36747;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3112.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.

```
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-451937/48.
XX
PT   Isolated polypeptide for treating, preventing and/ or prognosing
PT   disorders related to the musculoskeletal system including
PT   musculoskeletal cancers and also for testing and detection e.g.
PT   diagnosis -
XX
PS   Example 2; SEQ ID NO 3112; 781pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC   (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC   for preventing, treating or ameliorating medical conditions e.g. by
CC   protein or gene therapy. The genes are isolated from a range of human
CC   tissues disclosed in the specification. The nucleic acids, proteins,
CC   antibodies and (ant)agonists are useful in the diagnosis, treatment
CC   and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC   other cancers of the adrenal gland, bone, bone marrow, breast,
CC   gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC   (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC   healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC   and (f) infectious diseases such as viral, bacterial, fungal and
CC   parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5749 BP; 1600 A; 1192 C; 1403 G; 1553 T; 1 other;

  Query Match             100.0%;  Score 21;  DB 22;  Length 5749;
  Best Local Similarity   100.0%;  Pred. No. 2.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db     120 ACAAGACCTCAGACTTCCAGC 100


RESULT 3
AAT96690/c
ID   AAT96690 standard; DNA; 10825 BP.
XX
AC   AAT96690;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis;
KW   gene therapy; prenatal screening; human; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             361..7147
FT                   /*tag=  a
FT                   /note= "contains introns"
FT   intron          437..3761
FT                   /*tag=  b
FT                   /number= 1
FT   intron          4026..4234
FT                   /*tag=  c
FT                   /number= 2
FT   intron          4511..5605
FT                   /*tag=  d
FT                   /number= 3
FT   intron          5882..6039
FT                   /*tag=  e
FT                   /number= 4
FT   intron          6154..7106
FT                   /*tag=  f
FT                   /number= 5
FT   mutation        3872
FT                   /*tag=  g
FT                   /note= "C to G substitution (24d2 mutation)
FT                          results in His to Asp substitution"
FT   variation       3878
FT                   /*tag=  h
FT                   /note= "A to T substitution (24d7 variant)
FT                          results in Ser to Cys substitution"
FT   mutation        5834
FT                   /*tag=  i
FT                   /note= "G to A substitution (24d1 mutation
FT                          associated with HH), results in Cys to
FT                          Tyr substitution"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US06254.
XX
PR   23-MAY-1996;   96US-0652265.
PR   04-APR-1996;   96US-0630912.
```

PR 16-APR-1996; 96US-0632673.
XX
PA (MERC-) MERCATOR GENETICS INC.
XX
PI Drayna DT, Feder JN, Gnirke A, Ruddy D, Thomas WJ;
PI Tsuchihashi Z, Wolff RK;
XX
DR WPI; 1997-512743/47.
DR P-PSDB; AAW36499.
XX
PT Hereditary haemochromatosis gene and variants - useful for diagnosis
PT and treatment of hereditary haemochromatosis disease
XX
PS Disclosure; Fig 3; 115pp; English.
XX
CC This genomic DNA sequence corresponds to the human gene whose
CC mutated form is associated with hereditary haemochromatosis (HH).
CC To identify this novel gene, allelic association patterns were
CC determined between known markers and the HH locus in the HLA region
CC of chromosome 6. A physical clone coverage was then generated
CC extending from D6S265, which is a marker that is centromeric of
CC HLA-A, in a telomeric direction through D6S276, a marker at which
CC the allelic association was no longer observed. A single mutation
CC (24d1) in the HH gene appears responsible for the majority of HH
CC disease. This comprises a G to A substitution that is present in
CC 86% of affected chromosomes and in 4% of unaffected chromosomes.
CC It results in a Cys to Tyr substitution in the encoded protein (see
CC AAW36499) at a critical disulphide bridge important for secondary
CC structure. The following are claimed: the HH genomic DNA (I), a
CC 1437 bp cDNA sequence (Ia) (see AAT96691) and their 24d1, 24d2 and
CC 24d7 variants; a cloning or expression vector; host cells; a
CC peptide product chosen from the HH gene product, its variants
CC (24d1, 24d2 and 24d7), or a peptide of at least 56 amino acid
CC residues of these; an antibody produced using the peptide; a method
CC to determine the presence or absence of the common HH gene
CC mutation; an animal model for the HH disease; metal chelation
CC agents, T-cell differentiation factors and therapeutic agents for
CC the mitigation of injury due to oxidative process in vivo or
CC mitigation of iron overload; a method for screening potential
CC therapeutic agents for activity in connection with HH disease; an
CC antisense oligonucleotide directed against a transcriptional
CC product of a nucleic acid sequence as above; and oligonucleotides
CC or pairs of oligonucleotides covering a range of nucleotides from
CC (I), (Ia) or their variants, useful for detecting a polymorphism in
CC the HH gene. The invention also relates to methods for screening
CC for HH homozygotes, to HH diagnosis, prenatal screening and
CC diagnosis, and therapies of HH disease, including gene therapy,
CC protein- and antibody-based therapeutics, and small molecule
CC therapeutics.
XX
SQ Sequence 10825 BP; 2996 A; 2254 C; 2648 G; 2927 T; 0 other;

Query Match 100.0%; Score 21; DB 18; Length 10825;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
||||||||||||||||||||
Db 4120 ACAAGACCTCAGACTTCCAGC 4100

RESULT 4
AAC68425/c
ID AAC68425 standard; DNA; 10825 BP.
XX
AC AAC68425;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36869.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2998 A; 2253 C; 2648 G; 2926 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
||||||||||||||||||||
Db 4120 ACAAGACCTCAGACTTCCAGC 4100

RESULT 5
AAC68426/c
ID AAC68426 standard; DNA; 10825 BP.
XX
AC AAC68426;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1 mutation DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;

PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36870.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2999 A; 2253 C; 2647 G; 2926 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
||||||||||||||||||||
Db 4120 ACAAGACCTCAGACTTCCAGC 4100

RESULT 6
AAC68427/c
ID AAC68427 standard; DNA; 10825 BP.
XX
AC AAC68427;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d2 mutation DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36871.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2998 A; 2252 C; 2649 G; 2926 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
||||||||||||||||||||
Db 4120 ACAAGACCTCAGACTTCCAGC 4100

RESULT 7
AAC68428/c
ID AAC68428 standard; DNA; 10825 BP.
XX
AC AAC68428;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1/2 mutation DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36872.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2999 A; 2252 C; 2648 G; 2926 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
||||||||||||||||||||
Db 4120 ACAAGACCTCAGACTTCCAGC 4100

RESULT 8
AAA96794/c
ID AAA96794 standard; cDNA; 12146 BP.

```
XX
AC   AAA96794;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Genomic DNA of a histocompatibility iron loading (HFE) gene.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            1028..1324
FT                   /*tag= a
FT                   /number= 1
FT   intron          1325..4651
FT                   /*tag= b
FT                   /number= 1
FT   exon            4652..4915
FT                   /*tag= c
FT                   /number= 2
FT   intron          4916..5124
FT                   /*tag= d
FT                   /number= 2
FT   exon            5125..5400
FT                   /*tag= e
FT                   /number= 3
FT   intron          5401..6493
FT                   /*tag= f
FT                   /number= 3
FT   exon            6494..6769
FT                   /*tag= g
FT                   /number= 4
FT   intron          6770..6927
FT                   /*tag= h
FT                   /number= 4
FT   exon            6928..7041
FT                   /*tag= i
FT                   /number= 5
FT   intron          7042..7994
FT                   /*tag= j
FT                   /number= 5
FT   exon            7995..9050
FT                   /*tag= k
FT                   /number= 6
FT   intron          9051..10205
FT                   /*tag= l
FT                   /number= 6
FT   exon            10206..10637
FT                   /*tag= m
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Example 1; Page 21-28; 55pp; English.
XX
CC   The present sequence represents the human histocompatibility iron
CC   loading (HFE) gene. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence 12146 BP; 3383 A; 2474 C; 2911 G; 3378 T; 0 other;

  Query Match             100.0%;  Score 21;  DB 21;  Length 12146;
  Best Local Similarity   100.0%;  Pred. No. 2.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db    5010 ACAAGACCTCAGACTTCCAGC 4990


RESULT 9
AAV57926
ID   AAV57926 standard; DNA; 235033 BP.
XX
AC   AAV57926;
XX
DT   23-DEC-1998  (first entry)
XX
DE   Hereditary haemochromatosis subregion from an unaffected individual.
XX
KW   Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
KW   diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;
KW   BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
KW   type 1 sodium transport gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO9814466-A1.
XX
PD   09-APR-1998.
XX
PF   30-SEP-1997;   97WO-US17658.
XX
PR   07-MAY-1997;   97US-0852495.
PR   01-OCT-1996;   96US-0724394.
XX
PA   (PROG-) PROGENTIOR INC.
XX
PI   Feder JN,  Kronmal GS,  Lauer PM,  Ruddy DA,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1998-240014/21.
XX
PT   Hereditary haemochromatosis gene products - used to develop products
PT   for the diagnosis and treatment of hereditary disorders in iron
PT   metabolism
XX
PS   Example 2; Fig 8; 209pp; English.
XX
CC   The present invention describes hereditary haemochromatosis gene
CC   products from the human haemochromatosis gene. The present sequence
CC   represents a hereditary haemochromatosis subregion from an individual
CC   unaffected by hereditary haemochromatosis (HH). Also described is a
CC   method to determine the presence or absence of the common hereditary
CC   haemochromatosis (HFE) gene mutation in an individual comprising:
CC   (a) providing DNA or RNA from the individual; and (b) assessing the
CC   DNA or RNA for the presence or absence of a haplotype or genotype where
CC   the presence or absence of the haplotype genotype indicates the likely
```

CC presence of the HFE gene mutation in the genome of the individual. The
CC HFE gene sequences from the present invention can be used to develop
CC products for use in the diagnosis and treatment of HFE. The present
CC invention also describes BTF genes, which are homologues of the milk
CC protein butyrophilin (BT), and can be used in the production of agonists
CC and antagonists of BT function. Also described are: (1) a RoRet gene
CC which can be used to develop products for the study, diagnosis and
CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
CC which are homologues of a type 1 sodium transport gene, and can
CC similarly be used for hypophosphatemia.
XX
SQ Sequence 235033 BP; 68786 A; 48466 C; 49441 G; 68340 T; 0 other;

Query Match 100.0%; Score 21; DB 19; Length 235033;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
||||||||||||||||||||
Db 43030 ACAAGACCTCAGACTTCCAGC 43050

RESULT 10
AAV57903
ID AAV57903 standard; DNA; 237326 BP.
XX
AC AAV57903;
XX
DT 21-DEC-1998 (first entry)
XX
DE Hereditary haemochromatosis subregion from an HH affected individual.
XX
KW Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
KW diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;
KW BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
KW type 1 sodium transport gene; ss.
XX
OS Homo sapiens.
XX
PN WO9814466-A1.
XX
PD 09-APR-1998.
XX
PF 30-SEP-1997; 97WO-US17658.
XX
PR 07-MAY-1997; 97US-0852495.
PR 01-OCT-1996; 96US-0724394.
XX
PA (PROG-) PROGENTIOR INC.
XX
PI Feder JN, Kronmal GS, Lauer PM, Ruddy DA, Thomas WJ;
PI Tsuchihashi Z, Wolff RK;
XX
DR WPI; 1998-240014/21.
XX
PT Hereditary haemochromatosis gene products - used to develop products
PT for the diagnosis and treatment of hereditary disorders in iron
PT metabolism
XX
PS Claim 1; Fig 9; 209pp; English.
XX
CC The present invention describes hereditary haemochromatosis gene
CC products from the human haemochromatosis gene. The present sequence
CC represents a hereditary haemochromatosis subregion from an hereditary
CC haemochromatosis (HH) affected individual. Also described is a
CC method to determine the presence or absence of the common hereditary
CC haemochromatosis (HFE) gene mutation in an individual comprising:
CC (a) providing DNA or RNA from the individual; and (b) assessing the
CC DNA or RNA for the presence or absence of a haplotype or genotype where
CC the presence or absence of the haplotype genotype indicates the likely
CC presence of the HFE gene mutation in the genome of the individual. The
CC HFE gene sequences from the present invention can be used to develop
CC products for use in the diagnosis and treatment of HFE. The present
CC invention also describes BTF genes, which are homologues of the milk
CC protein butyrophilin (BT), and can be used in the production of agonists
CC and antagonists of BT function. Also described are: (1) a RoRet gene
CC which can be used to develop products for the study, diagnosis and
CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
CC which are homologues of a type 1 sodium transport gene, and can
CC similarly be used for hypophosphatemia.
XX
SQ Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;

Query Match 100.0%; Score 21; DB 19; Length 237326;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
||||||||||||||||||||
Db 42980 ACAAGACCTCAGACTTCCAGC 43000

RESULT 11
AAF11290
ID AAF11290 standard; cDNA; 624 BP.
XX
AC AAF11290;
XX
DT 13-MAR-2001 (first entry)
XX
DE Aspergillus niger EST SEQ ID NO:3813.
XX
KW Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Aspergillus niger.
XX
PN WO200056762-A2.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000WO-US07781.
XX
PR 22-MAR-1999; 99US-0273623.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO NORDISK AS.
XX
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX
DR WPI; 2000-594572/56.
XX
PT Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags -
XX
PS Claim 87; Page 1708-1709; 3161pp; English.
XX
CC The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring
CC the global expression of genes from FF cells allows the production
CC potential of the microorganisms to be improved. New genes may be
CC discovered, possible functions of unknown open reading frames can be
CC identified and gene copy number variation and stability can be
CC monitored. The expression of genes can be used to study how FF cells
CC adapt to changes in culture conditions, environmental stress, spore
CC morphogenesis, recombination, metabolic or catabolic pathway

```
CC   engineering. Using ESTs provides several advantages over genomic or
CC   random cDNA clones including elimination of redundancy as one spot on an
CC   array equals one gene or open reading frame, and organisation of the
CC   microarrays based on function of the gene products to facilitate
CC   analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC   Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
CC   niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
CC   AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
CC   all specifically claimed in the present invention.
XX
SQ   Sequence 624 BP; 138 A; 183 C; 149 G; 139 T; 15 other;

  Query Match            87.6%;  Score 18.4;  DB 21;  Length 624;
  Best Local Similarity  95.0%;  Pred. No. 29;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAG 20
          |||||||||| ||||||||
Db    197 ACAAGACCTCTGACTTCCAG 216


RESULT 12
AAK54198
ID   AAK54198 standard; cDNA; 440 BP.
XX
AC   AAK54198;
XX
DT   16-NOV-2001  (first entry)
XX
DE   Murine transport and binding associated protein encoding cDNA SEQ ID 763.
XX
KW   Murine; liver; gene library; amino acid synthesis; binding protein;
KW   cell metabolism; energy metabolism; fatty acid metabolism; synthesis;
KW   phospholipid metabolism; purine; pyrimidine; nucleoside; nucleotide;
KW   replication; transcription; translation; transport protein; ss.
XX
OS   Mus musculus.
XX
PN   DE20103510-U1.
XX
PD   07-JUN-2001.
XX
PF   28-FEB-2001; 2001DE-2003510.
XX
PR   02-DEC-1999;.  99DE-1058160.
XX
PA   (LION-) LION BIOSCIENCE AG.
XX
DR   WPI; 2001-368570/39.
XX
PT   Gene library containing sequences with specific 3'-ends and no polyA
PT   tail, encoding proteins involved in a wide range of cellular processes
PT   -
XX
PS   Claim 15; Page 230-231; 251pp; German.
XX
CC   This invention describes a novel gene library (A) comprises a gene
CC   sequence (or its part) encoding a protein involved in amino acid
CC   synthesis, cellular/energy metabolism, metabolism of
CC   fatty acids/phospholipids, synthesis or breakdown of
CC   purines/pyrimidines/nucleosides/nucleotides, DNA
CC   replication/transcription/translation, or is a transport/binding protein.
CC   (A) are produced that correspond to the 3'-end of mRNA but without the
CC   polyA tail. They can be prepared more efficiently and with less effort
CC   than conventional libraries. AAK53436-AAK54275 represent fragments of the
CC   gene library described in the method of the invention.
XX
SQ   Sequence 440 BP; 109 A; 135 C; 83 G; 113 T; 0 other;

  Query Match            84.8%;  Score 17.8;  DB 22;  Length 440;
  Best Local Similarity  90.5%;  Pred. No. 54;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          | |||| ||||||||||||||
Db    260 AAAAGAGCTCAGACTTCCAGC 280


RESULT 13
ABA43533/c
ID   ABA43533 standard; DNA; 485 BP.
XX
AC   ABA43533;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #2228.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes,
PT   useful for measuring gene expression in sample derived from human
PT   breast, comprises number of single exon nucleic acid probes  -
XX
PS   Claim 1; SEQ ID NO 2228; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for
CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far greater
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 485 BP; 101 A; 132 C; 104 G; 148 T; 0 other;

  Query Match            84.8%;  Score 17.8;  DB 22;  Length 485;
  Best Local Similarity  90.5%;  Pred. No. 54;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          ||  ||||||||||||||||
```

```
Db      79 ACTTGACCTCAGACTTCCAGC 59


RESULT 14
ABA53984/c
ID   ABA53984 standard; DNA; 485 BP.
XX
AC   ABA53984;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #2289.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 1; SEQ ID NO 2289; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 485 BP; 101 A; 132 C; 104 G; 148 T; 0 other;

  Query Match             84.8%;  Score 17.8;  DB 22;  Length 485;
  Best Local Similarity   90.5%;  Pred. No. 54;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           ||  ||||||||||||||||||
Db      79 ACTTGACCTCAGACTTCCAGC 59


RESULT 15
ABA23734/c
ID   ABA23734 standard; DNA; 485 BP.
XX
AC   ABA23734;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #2200 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 1; SEQ ID No 2200; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 485 BP; 101 A; 132 C; 104 G; 148 T; 0 other;

  Query Match             84.8%;  Score 17.8;  DB 22;  Length 485;
  Best Local Similarity   90.5%;  Pred. No. 54;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           ||  ||||||||||||||||||
Db      79 ACTTGACCTCAGACTTCCAGC 59


Search completed: November  2, 2002, 04:20:41
Job time : 48.2775 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:55:33 ; Search time 0.832418 Seconds
(without alignments)
6196.774 Million cell updates/sec

Title: US-09-981-606-16
Perfect score: 21
Sequence: 1 acaagacctcagacttccagc 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		21	100.0	21	4	US-09-277-457-16	Sequence 16, Appl
2	c	21	100.0	10825	3	US-08-652-265-1	Sequence 1, Appli
3	c	21	100.0	10825	3	US-08-652-265-3	Sequence 3, Appli
4	c	21	100.0	10825	3	US-08-652-265-5	Sequence 5, Appli
5	c	21	100.0	10825	3	US-08-652-265-7	Sequence 7, Appli
6	c	21	100.0	10825	3	US-08-834-497A-1	Sequence 1, Appli
7	c	21	100.0	10825	3	US-08-834-497A-3	Sequence 3, Appli
8	c	21	100.0	10825	3	US-08-834-497A-5	Sequence 5, Appli
9	c	21	100.0	10825	3	US-08-834-497A-7	Sequence 7, Appli
10	c	21	100.0	10825	4	US-09-503-444A-1	Sequence 1, Appli
11	c	21	100.0	10825	4	US-09-503-444A-3	Sequence 3, Appli
12	c	21	100.0	10825	4	US-09-503-444A-5	Sequence 5, Appli
13	c	21	100.0	10825	4	US-09-503-444A-7	Sequence 7, Appli
14	c	21	100.0	12146	4	US-09-277-457-27	Sequence 27, Appl
15	c	21	100.0	246240	2	US-08-724-394A-20	Sequence 20, Appl
16	c	21	100.0	246240	2	US-08-724-394A-21	Sequence 21, Appl
17	c	21	100.0	246240	2	US-08-724-394A-22	Sequence 22, Appl
18	c	16.4	78.1	3249	1	US-08-106-493A-1	Sequence 1, Appli
19	c	16.4	78.1	3249	1	US-08-429-264-1	Sequence 1, Appli
20	c	16.4	78.1	4853	1	US-08-832-883-1	Sequence 1, Appli
21	c	16.4	78.1	4853	2	US-08-832-877-1	Sequence 1, Appli
22	c	16.2	77.1	1215	1	US-08-592-214A-1	Sequence 1, Appli
23	c	16.2	77.1	1215	3	US-09-149-976-1	Sequence 1, Appli
24	c	16.2	77.1	1705	4	US-08-702-665A-2	Sequence 2, Appli
25	c	16.2	77.1	1714	4	US-09-151-102-3	Sequence 3, Appli
26	c	16.2	77.1	1714	4	US-08-929-846-3	Sequence 3, Appli
27		15.8	75.2	12225	2	US-08-822-445-11	Sequence 11, Appl
28		15.8	75.2	12225	4	US-09-396-540-11	Sequence 11, Appl
29		15.8	75.2	12616	2	US-08-822-445-9	Sequence 9, Appli
30		15.8	75.2	12616	4	US-09-396-540-9	Sequence 9, Appli
31		15.2	72.4	846	2	US-08-619-542B-43	Sequence 43, Appl
32		15.2	72.4	2972	2	US-08-720-484A-3	Sequence 3, Appli
33		15.2	72.4	2972	3	US-08-953-823A-3	Sequence 3, Appli
34	c	15.2	72.4	4403765	4	US-09-103-840A-2	Sequence 2, Appli
35	c	14.8	70.5	1909	3	US-09-100-193-6	Sequence 6, Appli
36	c	14.8	70.5	2171	3	US-08-851-843A-100	Sequence 100, App
37	c	14.8	70.5	2171	4	US-08-974-549A-266	Sequence 266, App
38	c	14.8	70.5	2171	4	US-08-854-050-100	Sequence 100, App
39	c	14.8	70.5	2171	4	US-09-430-323-100	Sequence 100, App
40	c	14.8	70.5	2176	4	US-08-974-549A-3	Sequence 3, Appli
41	c	14.8	70.5	3855	4	US-08-974-549A-4	Sequence 4, Appli
42	c	14.8	70.5	4015	3	US-08-851-843A-224	Sequence 224, App
43	c	14.8	70.5	4015	4	US-08-974-549A-1	Sequence 1, Appli
44	c	14.8	70.5	4015	4	US-08-854-050-224	Sequence 224, App
45	c	14.8	70.5	4015	4	US-09-430-323-224	Sequence 224, App

ALIGNMENTS

```
RESULT 1
US-09-277-457-16
; Sequence 16, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg, Barry E.
;  APPLICANT:  Sawada-Hirai, Ritsuko
;  APPLICANT:  Barton, James C.
;  TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
;  FILE REFERENCE: 10653/002001
;  CURRENT APPLICATION NUMBER: US/09/277,457
;  CURRENT FILING DATE:  1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Reverse Primer
US-09-277-457-16

  Query Match             100.0%;  Score 21;  DB 4;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 0.051;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db       1 ACAAGACCTCAGACTTCCAGC 21


RESULT 2
US-08-652-265-1/c
; Sequence 1, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
```

```
;       COUNTRY:  USA
;       ZIP:  94111-3834
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/652,265
;       FILING DATE:  23-MAY-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Smith, William M.
;       REGISTRATION NUMBER:  30,223
;       REFERENCE/DOCKET NUMBER:  17957-000500
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 576-0200
;       TELEFAX:  (415) 576-0300
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10825 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;       LOCATION:  6040..6153, 7107..7147)
;       OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /note= "No. 6025130mal or wild-type (unaffected)
;       OTHER INFORMATION:  Hereditary Hemochromatosis (HH) gene
;       OTHER INFORMATION:  allele"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  140..7319
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) allele
;       OTHER INFORMATION:  cDNA (SEQ ID NO:9)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  3852..3891
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d2(C)
;       OTHER INFORMATION:  allele (SEQ ID NO:41)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  5507..6023
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d1(G)
;       OTHER INFORMATION:  allele (SEQ ID NO:20)"
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3872, "c")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3878, "a")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d7
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "g")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d1
US-08-652-265-1
```

```
  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db    4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 3
US-08-652-265-3/c
; Sequence 3, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d1 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d1 allele cDNA (SEQ ID NO:10)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
```

```
;      OTHER INFORMATION:  for 24d2(C) allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-3

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db   4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 4
US-08-652-265-5/c
; Sequence 5, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d2 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d2 allele cDNA (SEQ ID NO:11)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(G) allele (SEQ ID NO:20)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-08-652-265-5

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db   4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 5
US-08-652-265-7/c
; Sequence 7, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  and 24d2 mutations"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  cDNA containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles
;      OTHER INFORMATION:  (SEQ ID NO:12)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-7

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db    4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 6
US-08-834-497A-1/c
; Sequence 1, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /note= "No. 6140305mal or wild-type (unaffected)
;      OTHER INFORMATION:  Hereditary Hemochromatosis (HH) gene
;      OTHER INFORMATION:  allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) allele
;      OTHER INFORMATION:  cDNA (SEQ ID NO:9)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;      OTHER INFORMATION:  sequence surrounding variant for 24d2(C)
;      OTHER INFORMATION:  allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;      OTHER INFORMATION:  sequence surrounding variant for 24d1(G)
```

```
;       OTHER INFORMATION:  allele (SEQ ID NO:20)"
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3872, "c")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3878, "a")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d7
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "g")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d1
US-08-834-497A-1

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db   4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 7
US-08-834-497A-3/c
; Sequence 3, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d1 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d1 allele cDNA (SEQ ID NO:10)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(C) allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-834-497A-3

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          |||||||||||||||||||||
Db   4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 8
US-08-834-497A-5/c
; Sequence 5, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
```

```
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d2 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d2 allele cDNA (SEQ ID NO:11)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(G) allele (SEQ ID NO:20)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
```

```
US-08-834-497A-5

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db    4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 9
US-08-834-497A-7/c
; Sequence 7, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
```

```
;       OTHER INFORMATION:
;       OTHER INFORMATION:  and 24d2 mutations"
;       OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;       OTHER INFORMATION:  gene containing a combination of both
;       OTHER INFORMATION:  24d1 and 24d2 alleles"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  140..7319
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  cDNA containing a combination of both
;       OTHER INFORMATION:  24d1 and 24d2 alleles
;       OTHER INFORMATION:  (SEQ ID NO:12)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  3852..3891
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  genomic sequence surrounding variant
;       OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  5507..6023
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  genomic sequence surrounding variant
;       OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3872, "g")
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "a")
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d1
US-08-834-497A-7

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db    4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 10
US-09-503-444A-1/c
; Sequence 1, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;     APPLICANT:  Thomas, Winston J.
;     APPLICANT:  Drayna, Dennis T.
;     APPLICANT:  Feder, John N.
;     APPLICANT:  Gnirke, Andreas
;     APPLICANT:  Ruddy, David
;     APPLICANT:  Tsuchihashi, Zenta
;     APPLICANT:  Wolff, Roger K.
;     TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;     NUMBER OF SEQUENCES:  44
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Pennie & Edmonds LLP
;       STREET:  1155 Avenue of the Americas
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows 95
;       SOFTWARE:  WordPerfect Version 8
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/503,444A
;       FILING DATE:  14-Feb-2000
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/652,265
;       FILING DATE:  23-May-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/632,673
;       FILING DATE:  16-Apr-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/630,912
;       FILING DATE:  04-Apr-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Poissant, Brian M.
;       REGISTRATION NUMBER:  28,462
;       REFERENCE/DOCKET NUMBER:  8907-0088-999
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-790-9090
;       TELEFAX:  212-869-9741
;       TELEX:  66141
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10825 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;       LOCATION:  6040..6153, 7107..7147)
;       OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /note= "No. 6228594mal or wild-type (unaffected)
;       OTHER INFORMATION:  Hereditary Hemochromatosis (HH) gene
;       OTHER INFORMATION:  allele"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  140..7319
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) allele
;       OTHER INFORMATION:  cDNA (SEQ ID NO:9)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  3852..3891
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d2(C)
;       OTHER INFORMATION:  allele (SEQ ID NO:41)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  5507..6023
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d1(G)
;       OTHER INFORMATION:  allele (SEQ ID NO:20)"
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3872, "c")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3878, "a")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d7
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "g")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
```

```
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-1

  Query Match             100.0%;  Score 21;  DB 4;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db    4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 11
US-09-503-444A-3/c
; Sequence 3, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d1 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d1 allele cDNA (SEQ ID NO:10)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(C) allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-3

  Query Match             100.0%;  Score 21;  DB 4;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db    4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 12
US-09-503-444A-5/c
; Sequence 5, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d2 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d2 allele cDNA (SEQ ID NO:11)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(G) allele (SEQ ID NO:20)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-09-503-444A-5

  Query Match             100.0%;  Score 21;  DB 4;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db    4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 13
US-09-503-444A-7/c
; Sequence 7, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  and 24d2 mutations"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  cDNA containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles
;      OTHER INFORMATION:  (SEQ ID NO:12)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
```

```
;       LOCATION:  replace(3872, "g")
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "a")
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d1
US-09-503-444A-7

  Query Match             100.0%;  Score 21;  DB 4;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db    4120 ACAAGACCTCAGACTTCCAGC 4100


RESULT 14
US-09-277-457-27/c
; Sequence 27, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg, Barry E.
;  APPLICANT:  Sawada-Hirai, Ritsuko
;  APPLICANT:  Barton, James C.
;  TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
;  FILE REFERENCE: 10653/002001
;  CURRENT APPLICATION NUMBER: US/09/277,457
;  CURRENT FILING DATE:  1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
;   LENGTH: 12146
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-09-277-457-27

  Query Match             100.0%;  Score 21;  DB 4;  Length 12146;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db    5010 ACAAGACCTCAGACTTCCAGC 4990


RESULT 15
US-08-724-394A-20/c
; Sequence 20, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;     APPLICANT:  Feder, John N.
;     APPLICANT:  Kronmal, Gregory S.
;     APPLICANT:  Lauer, Peter M.
;     APPLICANT:  Ruddy, David A.
;     APPLICANT:  Thomas, Winston
;     APPLICANT:  Tsuchihashi, Zenta
;     APPLICANT:  Wolff, Roger K.
;     TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;     TITLE OF INVENTION:  Sequences and Antibodies Thereto
;     NUMBER OF SEQUENCES:  31
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;       STREET:  Two Embarcadero Center, 8th Floor
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94111-3834
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/724,394A
;       FILING DATE:  01-OCT-1996
;       CLASSIFICATION:  536
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Fitts, Renee A.
;       REGISTRATION NUMBER:  35,136
;       REFERENCE/DOCKET NUMBER:  017957-000100
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-576-0200
;       TELEFAX:  415-576-0300
;   INFORMATION FOR SEQ ID NO:  20:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  246240 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  not relevant
;       TOPOLOGY:  not relevant
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  1..246240
;       OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-20

  Query Match             100.0%;  Score 21;  DB 2;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           |||||||||||||||||||||
Db  196423 ACAAGACCTCAGACTTCCAGC 196403


Search completed: November  2, 2002, 06:45:35
Job time : 35.8324 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:10:18 ; Search time 31.772 Seconds
(without alignments)
8920.945 Million cell updates/sec

Title: US-09-981-606-16
Perfect score: 21
Sequence: 1 acaagacctcagacttccagc 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	18.4	87.6	767	10	BI769743	BI769743 603055044
c 2	17.8	84.8	342	9	AV683034	AV683034 AV683034
c 3	17.8	84.8	359	9	AA444544	AA444544 vf79e11.r
c 4	17.8	84.8	378	9	BB808342	BB808342 BB808342
c 5	17.8	84.8	399	9	AV693380	AV693380 AV693380
c 6	17.8	84.8	404	9	BB797456	BB797456 BB797456
c 7	17.8	84.8	407	12	AZ284463	AZ284463 RPCI-23-1
8	17.8	84.8	417	10	BG793419	BG793419 UTSW_SM11
c 9	17.8	84.8	447	9	AA980775	AA980775 ua45e02.r
c 10	17.8	84.8	461	9	AV734807	AV734807 AV734807
11	17.8	84.8	488	9	AW411615	AW411615 uq45c06.x
c 12	17.8	84.8	496	9	AI183057	AI183057 ub93c11.r
13	17.8	84.8	506	9	AW240817	AW240817 uq34g05.x
14	17.8	84.8	515	9	AW464712	AW464712 BP230016B
15	17.8	84.8	522	10	BE852898	BE852898 uw35g12.x
c 16	17.8	84.8	639	9	AW258421	AW258421 uq34g05.y
c 17	17.8	84.8	640	9	AV687303	AV687303 AV687303
c 18	17.8	84.8	661	9	AV834360	AV834360 AV834360
19	17.8	84.8	681	10	BG907157	BG907157 TaLr1157F
20	17.8	84.8	700	9	AL503911	AL503911 AL503911
c 21	17.8	84.8	708	10	BI105955	BI105955 602890908
22	17.8	84.8	722	10	BG142046	BG142046 ia87f02.x
c 23	17.8	84.8	764	9	AA790263	AA790263 vw16c04.r
c 24	17.8	84.8	821	10	BI455976	BI455976 603170420
25	17.8	84.8	892	10	BE413301	BE413301 MCG012.G0
26	17.4	82.9	305	9	AW416731	AW416731 52328 MAR
27	17	81.0	444	10	BE684391	BE684391 184781 MA
c 28	17	81.0	448	12	AQ452086	AQ452086 HS_5064_B
29	17	81.0	453	12	AQ089603	AQ089603 HS_3008_A
c 30	17	81.0	475	12	AZ725985	AZ725985 RPCI-24-6
31	17	81.0	494	12	AQ201625	AQ201625 RPCI11-67
c 32	17	81.0	519	12	AQ177779	AQ177779 HS_2211_A
33	17	81.0	525	9	AW660200	AW660200 98113 MAR
34	17	81.0	533	12	AQ569586	AQ569586 HS_5346_B
35	17	81.0	556	10	BF398547	BF398547 UI-R-BS2-
36	17	81.0	636	12	AG115999	AG115999 Pan trogl
37	17	81.0	655	12	AQ732669	AQ732669 HS_5546_B
c 38	17	81.0	671	12	AG158321	AG158321 Pan trogl
39	17	81.0	672	12	AG158183	AG158183 Pan trogl
c 40	17	81.0	895	10	BI766128	BI766128 603052929
c 41	16.8	80.0	171	10	F00380	F00380 HSB32E111 S
42	16.8	80.0	179	10	Z19482	Z19482 HSB32E112 S
c 43	16.8	80.0	192	10	BF755343	BF755343 MR2-CT059
c 44	16.8	80.0	252	10	BI718089	BI718089 1031023G1
c 45	16.8	80.0	266	9	AW534627	AW534627 UI-R-C4-a

ALIGNMENTS

RESULT 1
BI769743/c
LOCUS BI769743 767 bp mRNA linear EST 25-SEP-2001
DEFINITION 603055044F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5204410 5', mRNA sequence.
ACCESSION BI769743
VERSION BI769743.1 GI:15761308
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 767)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11512 row: h column: 11
High quality sequence stop: 749.
FEATURES Location/Qualifiers
source 1. .767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5204410"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size

range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."
BASE COUNT 229 a 147 c 181 g 210 t
ORIGIN

Query Match 87.6%; Score 18.4; DB 10; Length 767;
Best Local Similarity 95.0%; Pred. No. 3.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGACCTCAGACTTCCAGC 21
||||||| |||||||||||
Db 186 CAAGACCGCAGACTTCCAGC 167

RESULT 2
AV683034/c
LOCUS AV683034 342 bp mRNA linear EST 16-JAN-2002
DEFINITION AV683034 GKC Homo sapiens cDNA clone GKCGHH07 5', mRNA sequence.
ACCESSION AV683034
VERSION AV683034.1 GI:10284897
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 342)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES Location/Qualifiers
source 1. .342
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCGHH07"
/clone_lib="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 69 a 55 c 94 g 124 t
ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 342;
Best Local Similarity 90.5%; Pred. No. 4.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
||||| ||| |||||||||||
Db 184 ACAAGGCCTAAGACTTCCAGC 164

RESULT 3
AA444544/c
LOCUS AA444544 359 bp mRNA linear EST 03-JUN-1997
DEFINITION vf79e11.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone IMAGE:850028 5' similar to gb:X74953 M.musculus ETL-2 mRNA (MOUSE
);; mRNA sequence.
ACCESSION AA444544
VERSION AA444544.1 GI:2157484
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 359)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:502180
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 354.
FEATURES Location/Qualifiers
source 1. .359
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:850028"
/clone_lib="Soares_mammary_gland_NbMMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 78 a 89 c 109 g 83 t
ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 359;
Best Local Similarity 90.5%; Pred. No. 4.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
| |||| |||||||||||||
Db 324 AAAAGAGCTCAGACTTCCAGC 304

RESULT 4
BB808342/c
LOCUS BB808342 378 bp mRNA linear EST 19-NOV-2001
DEFINITION BB808342 RIKEN full-length enriched, 16 days neonate male medulla oblongata Mus musculus cDNA clone G630087G17 3', mRNA sequence.
ACCESSION BB808342
VERSION BB808342.1 GI:16980971
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 378)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES                Location/Qualifiers
     source             1. .378
                        /organism="Mus musculus"
                        /strain="C57BL/6J"
                        /db_xref="taxon:10090"
                        /clone="G630087G17"
                        /clone_lib="RIKEN full-length enriched, 16 days neonate
                        male medulla oblongata"
                        /sex="male"
                        /tissue_type="medulla oblongata"
                        /dev_stage="16 days neonate"
BASE COUNT       84 a     79 c    114 g    101 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 378;
  Best Local Similarity   90.5%;  Pred. No. 4.9e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          | |||| ||||||||||||||
Db    132 AAAAGAGCTCAGACTTCCAGC 112


RESULT 5
AV693380/c
LOCUS       AV693380                399 bp    mRNA    linear   EST 16-JAN-2002
DEFINITION  AV693380 GKC Homo sapiens cDNA clone GKCDLA05 5', mRNA sequence.
ACCESSION   AV693380
VERSION     AV693380.1  GI:10295243
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.
  TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
  MEDLINE   21625106
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES                Location/Qualifiers
     source             1. .399
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="GKCDLA05"
                        /clone_lib="GKC"
                        /tissue_type="hepatocellular carcinoma"
                        /dev_stage="Adult"
                        /lab_host="SOLR"
                        /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                        XhoI"
BASE COUNT       86 a     75 c    114 g    124 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 399;
  Best Local Similarity   90.5%;  Pred. No. 5e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          ||||| ||| ||||||||||||
Db    298 ACAAGGCCTAAGACTTCCAGC 278


RESULT 6
BB797456/c
LOCUS       BB797456                404 bp    mRNA    linear   EST 16-NOV-2001
DEFINITION  BB797456 RIKEN full-length enriched, 16 days neonate male medulla
            oblongata Mus musculus cDNA clone G630006B19 3', mRNA sequence.
ACCESSION   BB797456
VERSION     BB797456.1  GI:16967075
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 404)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
```

```
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .404
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="G630006B19"
                     /clone_lib="RIKEN full-length enriched, 16 days neonate
                     male medulla oblongata"
                     /sex="male"
                     /tissue_type="medulla oblongata"
                     /dev_stage="16 days neonate"
BASE COUNT       94 a     76 c    125 g    109 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 404;
  Best Local Similarity   90.5%;  Pred. No. 5e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           | ||||  ||||||||||||||
Db     158 AAAAGAGCTCAGACTTCCAGC 138


RESULT 7
AZ284463/c
LOCUS       AZ284463                407 bp    DNA     linear   GSS 27-JUL-2000
DEFINITION  RPCI-23-125H5.TV RPCI-23 Mus musculus genomic clone RPCI-23-125H5,
            DNA sequence.
ACCESSION   AZ284463
VERSION     AZ284463.1  GI:9526249
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 407)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-125H5.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 125  row: H  column: 5
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .407
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-125H5"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).   "
BASE COUNT       86 a     68 c    130 g    123 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 12;  Length 407;
  Best Local Similarity   90.5%;  Pred. No. 5e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ACAAGACCTCAGACTTCCAGC 21
           | ||||  ||||||||||||||
Db     118 AAAAGAGCTCAGACTTCCAGC 98


RESULT 8
BG793419
LOCUS       BG793419                417 bp    mRNA    linear   EST 16-MAY-2001
DEFINITION  UTSW_SM11H9 UTSW Adult Mouse Skeletal Muscle Library Mus musculus
            cDNA clone UTSW_SM11H9, mRNA sequence.
ACCESSION   BG793419
VERSION     BG793419.1  GI:14128989
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 417)
  AUTHORS   Gallardo,T.D., Schageman,J.J., Pertsemlidis,A., Garner,H.R.,
            Williams,R.S. and Shohet,R.V.
  TITLE     UT Southwestern Medical Center, Adult Mouse Skeletal Muscle cDNA
            Library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Schageman JJ
            Shohet/Garner Labs
            University of Texas Southwestern Medical Center
            6000 Harry Hines Blvd., NA2.226, Dallas, TX 75390, USA
            Tel: 214 648 1674
            Email: Jeff.Schageman@UTSouthwestern.edu
            cDNA library constructed by UTSW as a component of the Program for
            Genomic Applications (PGA) and the Reynolds Heart Disease
            Prevention grants for use in cDNA microarray experiments. Sequence
            Quality: Sequence ends were trimmed based on percentage of ambigu
            us base calls or 'N's in windowed segments. Sequencing: First-pass
            sequencing; ABI Prism 377 sequencer and analysis software.
            Seq primer: M13/pUC Reverse.
FEATURES             Location/Qualifiers
     source          1. .417
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="UTSW_SM11H9"
```

/clone_lib="UTSW Adult Mouse Skeletal Muscle Library"
/sex="Pooled"
/tissue_type="Diaphragm/Hind limb muscles"
/cell_type="Skeletal muscle"
/dev_stage="2 months"
/lab_host="DH5a"
/note="Vector: pAMP10 (Gibco); Cloned unidirectionally. Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps (Manniatis); Cloning Technique: CUA Cloning (CloneAmp, Life Technologies); Average insert size: 1.8 Kb; Insertion site: TACGTCCACTGAATTCTGAGTG--->. Other information regarding entire library may be found at http://pga.swmed.edu/Data/Libraries/microarray_cdna_libraries.htm."

BASE COUNT 110 a 130 c 75 g 100 t 2 others
ORIGIN

Query Match 84.8%; Score 17.8; DB 10; Length 417;
Best Local Similarity 90.5%; Pred. No. 5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
| |||| |||||||||||||||
Db 288 AAAAGAGCTCAGACTTCCAGC 308

RESULT 9
AA980775/c
LOCUS AA980775 447 bp mRNA linear EST 27-MAY-1998
DEFINITION ua45e02.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone IMAGE:1349690 5' similar to gb:X74953 M.musculus ETL-2 mRNA (MOUSE);; mRNA sequence.
ACCESSION AA980775
VERSION AA980775.1 GI:3159311
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 447)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:698482
Putative full length read
vector to vector length is 449
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 364.
FEATURES Location/Qualifiers
source 1. .447
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1349690"
/clone_lib="Soares_mammary_gland_NbMMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 104 a 89 c 140 g 114 t
ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 447;
Best Local Similarity 90.5%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
| |||| |||||||||||||||
Db 206 AAAAGAGCTCAGACTTCCAGC 186

RESULT 10
AV734807/c
LOCUS AV734807 461 bp mRNA linear EST 17-OCT-2000
DEFINITION AV734807 cdA Homo sapiens cDNA clone cdAAOF05 5', mRNA sequence.
ACCESSION AV734807
VERSION AV734807.1 GI:10852352
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 461)
AUTHORS Yang,Y., Song,H., Peng,Y., Gu,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu ,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Chen,J., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA cdA clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES Location/Qualifiers
source 1. .461
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="cdAAOF05"
/clone_lib="cdA"
/tissue_type="pheochromocytoma"
/dev_stage="Adult"
/lab_host="BM25.8"
/note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT 110 a 81 c 129 g 140 t 1 others
ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 461;
Best Local Similarity 90.5%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACAAGACCTCAGACTTCCAGC 21
||||| ||| ||||||||||
Db 360 ACAAGGCCTAAGACTTCCAGC 340

RESULT 11
AW411615
LOCUS AW411615 488 bp mRNA linear EST 08-FEB-2000

```
DEFINITION  uq45c06.x1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:2812330 3'
            similar to gb:X74953 M.musculus ETL-2 mRNA (MOUSE);, mRNA sequence.
ACCESSION   AW411615
VERSION     AW411615.1  GI:6937470
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 488)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            MGI:1044942
            Seq primer: -40UP from Gibco
            High quality sequence stop: 359.
FEATURES             Location/Qualifiers
     source          1. .488
                     /organism="Mus musculus"
                     /strain="C57/B6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:2812330"
                     /clone_lib="NCI_CGAP_Mam5"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"
BASE COUNT      117 a    159 c     96 g    116 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 488;
  Best Local Similarity   90.5%;  Pred. No. 5.2e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          | ||||  ||||||||||||||
Db    244 AAAAGAGCTCAGACTTCCAGC 264


RESULT 12
AI183057/c
LOCUS       AI183057                496 bp    mRNA    linear   EST 08-OCT-1998
DEFINITION  ub93c11.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
            IMAGE:1396052 5' similar to gb:X74953 M.musculus ETL-2 mRNA (MOUSE
            );, mRNA sequence.
ACCESSION   AI183057
VERSION     AI183057.1  GI:3733695
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 496)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:907768
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 326.
FEATURES             Location/Qualifiers
     source          1. .496
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1396052"
                     /clone_lib="Soares_mammary_gland_NbMMG"
                     /sex="male"
                     /tissue_type="mammary gland"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                     constructed and normalized by Bento Soares and M.Fatima
                     Bonaldo."
BASE COUNT      100 a    100 c    151 g    145 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 496;
  Best Local Similarity   90.5%;  Pred. No. 5.2e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          | ||||  ||||||||||||||
Db    251 AAAAGAGCTCAGACTTCCAGC 231


RESULT 13
AW240817
LOCUS       AW240817                506 bp    mRNA    linear   EST 03-APR-2000
DEFINITION  uq34g05.x1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:2811320 3'
            similar to gb:X74953 M.musculus ETL-2 mRNA (MOUSE);, mRNA sequence.
ACCESSION   AW240817
VERSION     AW240817.1  GI:6574569
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 506)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
```

```
           www-bio.llnl.gov/bbrp/image/image.html
           Seq primer: -40UP from Gibco
           High quality sequence stop: 335.
FEATURES             Location/Qualifiers
     source          1. .506
                     /organism="Mus musculus"
                     /strain="C57/B6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:2811320"
                     /clone_lib="NCI_CGAP_Mam5"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"
BASE COUNT      124 a    161 c    102 g    119 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 506;
  Best Local Similarity   90.5%;  Pred. No. 5.2e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          | ||||  ||||||||||||||
Db    247 AAAAGAGCTCAGACTTCCAGC 267


RESULT 14
AW464712
LOCUS       AW464712                515 bp    mRNA    linear   EST 24-FEB-2000
DEFINITION  BP230016B10E12 Soares normalized bovine placenta Bos taurus cDNA
            clone BP230016B10E12 5', mRNA sequence.
ACCESSION   AW464712
VERSION     AW464712.1  GI:7034880
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 515)
  AUTHORS   Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,J., Liu,L. and Larson
            ,J.H.
  TITLE     Bovine ESTs
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Lewin, H. A.
            W. M. Keck Center for Comparative and Functional Genomics
            University of Illinois at Urbana-Champaign
            340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL
            61801, USA
            Tel: 217 333 5998
            Fax: 217 244 5617
            Email: h-lewin@uiuc.edu
            Funding for cattle EST sequencing was provided by the USDA National
            Research Initiative, Animal Genome Resource Grant AG 99-3205-8534
            to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
            from Washington University Genome Center. Vector Trimmi g:
            Cross_match from Washington University Genome Center PHRAP suite.
            Sequences submitted are vector free and at least 200 bp in length.
            PCR PRimers
            FORWARD: TAATACGACTCACTATAGGG
            BACKWARD: ATTAACCCTCACTAAAG
            Insert Length: 515   Std Error: 0.00
            Plate: BP230016B10  row: E  column: 12
            Seq primer: AGCGGATAACAATTTCACACAGGA
            High quality sequence stop: 515.
FEATURES             Location/Qualifiers
     source          1. .515
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone="BP230016B10E12"
                     /clone_lib="Soares normalized bovine placenta"
                     /sex="female"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pT7T3Pac; Site_1: EcoRI;
                     Site_2: NotI; The cDNA library was contributed by the
                     Soares laboratory and it  was constructed and normalized
                     as described by Bonaldo, M.F.,  Lennon, G. and Soares,
                     M.B. (1996), Genome Research 6(9): 791-806. "
BASE COUNT      144 a    115 c    126 g    129 t      1 others
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 515;
  Best Local Similarity   90.5%;  Pred. No. 5.2e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACAAGACCTCAGACTTCCAGC 21
          | ||||  ||||||||||||||
Db    224 ACAAGGCCGCAGACTTCCAGC 244


RESULT 15
BE852898
LOCUS       BE852898                522 bp    mRNA    linear   EST 26-SEP-2000
DEFINITION  uw35g12.x1 Soares_thymus_2NbMT Mus musculus cDNA clone
            IMAGE:3418726 3' similar to TR:P70225 P70225 INTERLEUKIN-11
            RECEPTOR ALPHA CHAIN 2 PRECURSOR ;, mRNA sequence.
ACCESSION   BE852898
VERSION     BE852898.1  GI:10311237
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Other_ESTs: uw35g12.y1
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1094538
            High quality sequence stop: 457.
FEATURES             Location/Qualifiers
     source          1. .522
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3418726"
                     /clone_lib="Soares_thymus_2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      129 a    166 c    102 g    125 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 10;  Length 522;
  Best Local Similarity   90.5%;  Pred. No. 5.2e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       1 ACAAGACCTCAGACTTCCAGC 21
           | |||| ||||||||||||||
Db     246 AAAAGAGCTCAGACTTCCAGC 266
```

Search completed: November 2, 2002, 06:42:16
Job time : 36.772 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:08:53 ; Search time 42.6923 Seconds
(without alignments)
10293.594 Million cell updates/sec

Title: US-09-981-606-4
Perfect score: 21
Sequence: 1 gctctgacaacctcaggaagg 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	21	100.0	249	9	HSU80914	U80914 Human hered
c 2	21	100.0	10825	6	AR117789	AR117789 Sequence
c 3	21	100.0	10825	6	AR117790	AR117790 Sequence
c 4	21	100.0	10825	6	AR117791	AR117791 Sequence
c 5	21	100.0	10825	6	AR117792	AR117792 Sequence
c 6	21	100.0	10825	6	AR149459	AR149459 Sequence
c 7	21	100.0	10825	6	AR149460	AR149460 Sequence
c 8	21	100.0	10825	6	AR149461	AR149461 Sequence
c 9	21	100.0	10825	6	AR149462	AR149462 Sequence
c 10	21	100.0	12146	9	HSHFE	Z92910 Homo sapien
11	21	100.0	193752	2	AL359892	AL359892 Homo sapi
c 12	21	100.0	246240	6	AR036572	AR036572 Sequence
c 13	21	100.0	246240	6	AR036573	AR036573 Sequence
c 14	21	100.0	246240	6	AR036574	AR036574 Sequence
15	21	100.0	246282	9	HSU91328	U91328 Human hered
c 16	18.4	87.6	874	9	HSHLAH2	Y09800 H.sapiens H
17	18.4	87.6	213565	2	AC096317	AC096317 Rattus no
18	18	85.7	141980	9	AL450425	AL450425 Human DNA
19	18	85.7	162299	2	AC016429	AC016429 Homo sapi
20	17.8	84.8	1138	8	AF411228	AF411228 Hordeum v
21	17.8	84.8	5995	1	ATUAUX	M61151 Agrobacteri
22	17.8	84.8	178867	2	AC074095	AC074095 Homo sapi
23	17.8	84.8	296950	1	AP001508	AP001508 Bacillus
c 24	17.8	84.8	314146	2	AC073759	AC073759 Mus muscu
25	17.8	84.8	337101	9	HSXKSRPXR	AL121578 Homo sapi
c 26	17.4	82.9	93418	9	AC008379	AC008379 Homo sapi
c 27	17.4	82.9	138851	2	AC078864	AC078864 Homo sapi
28	17.4	82.9	140416	2	AC093024	AC093024 Homo sapi
29	17.4	82.9	148654	2	AC018353	AC018353 Homo sapi
30	17.4	82.9	156795	2	AC022290	AC022290 Homo sapi
31	17.4	82.9	158285	9	AC021066	AC021066 Homo sapi
c 32	17.4	82.9	158305	2	AC023075	AC023075 Homo sapi
33	17.4	82.9	162810	9	AC006382	AC006382 Homo sapi
34	17.4	82.9	163277	2	AC084687	AC084687 Homo sapi
35	17.4	82.9	166046	2	AC080174	AC080174 Homo sapi
36	17.4	82.9	172272	2	AC090671	AC090671 Homo sapi
37	17.4	82.9	172716	2	AC105756	AC105756 Homo sapi
38	17.4	82.9	174256	2	AC022572	AC022572 Homo sapi
39	17.4	82.9	176749	9	AC096748	AC096748 Homo sapi
40	17.4	82.9	180885	2	AC107221	AC107221 Homo sapi
41	17.4	82.9	182084	2	AC026615	AC026615 Homo sapi
42	17.4	82.9	184037	9	CNS01RIM	AL163642 Human chr
43	17.4	82.9	208842	2	AC091295	AC091295 Mus muscu
c 44	17.4	82.9	320336	2	AC018352	AC018352 Homo sapi
c 45	17	81.0	81854	9	AC006396	AC006396 Homo sapi

ALIGNMENTS

RESULT 1
HSU80914/c
LOCUS HSU80914 249 bp DNA linear PRI 05-JAN-1999
DEFINITION Human hereditary haemochromatosis protein (HLA-H) gene, partial cds.
ACCESSION U80914
VERSION U80914.1 GI:4098856
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 249)
AUTHORS Hashimoto,K., Hirai,M. and Kurosawa,Y.
TITLE Identification of a mouse homolog for the human hereditary haemochromatosis candidate gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 249)
AUTHORS Hashimoto,K.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-1996) Institute for Comprehensive Medical Science, Fujita Health University, Aichi, Toyoake 470-11, Japan
FEATURES Location/Qualifiers

```
     source          1. .249
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22"
                     /tissue_type="placenta"
     mRNA            join(<1. .20,230. .>249)
                     /gene="HLA-H"
     gene            <1. .>249
                     /gene="HLA-H"
     CDS             join(<1. .20,230. .>249)
                     /gene="HLA-H"
                     /note="MR2(alpha 1/intron/alpha 2)"
                     /codon_start=2
                     /product="hereditary haemochromatosis protein"
                     /protein_id="AAD00449.1"
                     /db_xref="GI:4098857"
                     /translation="NHNHSKESHTLQV"
BASE COUNT       59 a     50 c     75 g     65 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 9;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db     62 GCTCTGACAACCTCAGGAAGG 42


RESULT 2
AR117789/c
LOCUS       AR117789               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6140305.
ACCESSION   AR117789
VERSION     AR117789.1  GI:14098695
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 1 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a    2253 c    2648 g    2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db   4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 3
AR117790/c
LOCUS       AR117790               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 3 from patent US 6140305.
ACCESSION   AR117790
VERSION     AR117790.1  GI:14098696
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 3 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a    2253 c    2647 g    2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db   4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 4
AR117791/c
LOCUS       AR117791               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 5 from patent US 6140305.
ACCESSION   AR117791
VERSION     AR117791.1  GI:14098697
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 5 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a    2252 c    2649 g    2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db   4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 5
AR117792/c
LOCUS       AR117792               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 7 from patent US 6140305.
ACCESSION   AR117792
VERSION     AR117792.1  GI:14098698
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 7 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a    2252 c    2648 g    2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
```

```
          ||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 6
AR149459/c
LOCUS       AR149459               10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 1 from patent US 6228594.
ACCESSION   AR149459
VERSION     AR149459.1  GI:15114050
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 1 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a   2253 c   2648 g   2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          ||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 7
AR149460/c
LOCUS       AR149460               10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 3 from patent US 6228594.
ACCESSION   AR149460
VERSION     AR149460.1  GI:15114051
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 3 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a   2253 c   2647 g   2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          ||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 8
AR149461/c
LOCUS       AR149461               10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 5 from patent US 6228594.
ACCESSION   AR149461
VERSION     AR149461.1  GI:15114052
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 5 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a   2252 c   2649 g   2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          ||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 9
AR149462/c
LOCUS       AR149462               10825 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 7 from patent US 6228594.
ACCESSION   AR149462
VERSION     AR149462.1  GI:15114053
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 7 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2999 a   2252 c   2648 g   2926 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          ||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 10
HSHFE/c
LOCUS       HSHFE                  12146 bp    DNA     linear   PRI 23-JUL-1999
DEFINITION  Homo sapiens HFE gene.
ACCESSION   Z92910
VERSION     Z92910.1  GI:1890179
KEYWORDS    haemochromatosis; HFE gene.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 858)
  AUTHORS   Albig,W., Drabent,B., Burmester,N., Bode,C. and Doenecke,D.
  TITLE     The haemochromatosis candidate gene HFE (HLA-H) of man and mouse is
            located in syntenic regions within the histone gene cluster
  JOURNAL   J. Cell. Biochem. 69 (2), 117-126 (1998)
  MEDLINE   98208340
REFERENCE   2  (bases 1 to 12146)
```

```
  AUTHORS   Albig,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-1997) Albig W., Georg-August-Universitaet
            Goettingen, Biochemie und Molekulare Zellbiologie, Humboldtallee
            23, Goettingen, FRG, 37073
FEATURES             Location/Qualifiers
     source          1. .12146
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p"
                     /clone="ICRFy901D1223"
                     /clone_lib="ICRF YAC-library"
     exon            1028. .1324
                     /gene="HFE"
                     /number=1
     gene            1028. .10637
                     /gene="HFE"
     CDS             join(1249. .1324,4652. .4915,5125. .5400,6494. .6769,
                     6928. .7041,7995. .8035)
                     /gene="HFE"
                     /function="iron metabolism"
                     /note="haemochromatosis candidate gene"
                     /codon_start=1
                     /protein_id="CAB07442.1"
                     /db_xref="GI:1890180"
                     /db_xref="SWISS-PROT:Q30201"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                     SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
                     WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                     RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTS
                     SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGE
                     EQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQG
                     SRGAMGHYVLAERE"
     intron          1325. .4651
                     /gene="HFE"
                     /number=1
     repeat_unit     3494. .3735
                     /gene="HFE"
                     /rpt_family="Alu"
     repeat_unit     3973. .4283
                     /gene="HFE"
                     /rpt_family="Alu"
     exon            4652. .4915
                     /gene="HFE"
                     /number=2
     intron          4916. .5124
                     /gene="HFE"
                     /number=2
     exon            5125. .5400
                     /gene="HFE"
                     /number=3
     intron          5401. .6493
                     /gene="HFE"
                     /number=3
     repeat_unit     5707. .6005
                     /gene="HFE"
                     /rpt_family="Alu"
     exon            6494. .6769
                     /gene="HFE"
                     /number=4
     intron          6770. .6927
                     /gene="HFE"
                     /number=4
     exon            6928. .7041
                     /gene="HFE"
                     /number=5
     intron          7042. .7994
                     /gene="HFE"
                     /number=5
     exon            7995. .9050
                     /gene="HFE"
                     /number=6
     repeat_unit     9017. .9340
                     /gene="HFE"
                     /rpt_family="Alu"
     intron          9051. .10205
                     /gene="HFE"
                     /number=6
     repeat_unit     9957. .10239
                     /gene="HFE"
                     /rpt_family="Alu"
     exon            10206. .10637
                     /gene="HFE"
                     /number=7
     polyA_signal    10617. .10622
                     /gene="HFE"
BASE COUNT     3383 a   2474 c   2911 g   3378 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 9;  Length 12146;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           |||||||||||||||||||||
Db    4957 GCTCTGACAACCTCAGGAAGG 4937


RESULT 11
AL359892
LOCUS       AL359892              193752 bp    DNA     linear   HTG 13-JUN-2001
DEFINITION  Homo sapiens chromosome 6 clone RP11-557F22, *** SEQUENCING IN
            PROGRESS ***, 18 unordered pieces.
ACCESSION   AL359892
VERSION     AL359892.5  GI:9930971
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 193752)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 27, 2000 this sequence version replaced gi:9864230.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bA557F22
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 183925 bases at least Q40
            Consensus quality: 187703 bases at least Q30
            Consensus quality: 189658 bases at least Q20
            Insert size: 192052; sum-of-contigs
            Insert size: 198247; agarose-fp
            Quality coverage: 3.68x in Q20 bases; sum-of-contigs Quality
            coverage: 3.70x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 18 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3250: contig of 3250 bp in length
```

```
            *     3251 3350: gap of      100 bp
            *     3351    14600: contig of 11250 bp in length
            *    14601 14700: gap of      100 bp
            *    14701    32357: contig of 17657 bp in length
            *    32358 32457: gap of      100 bp
            *    32458    34886: contig of 2429 bp in length
            *    34887 34986: gap of      100 bp
            *    34987    43490: contig of 8504 bp in length
            *    43491 43590: gap of      100 bp
            *    43591    47437: contig of 3847 bp in length
            *    47438 47537: gap of      100 bp
            *    47538    57356: contig of 9819 bp in length
            *    57357 57456: gap of      100 bp
            *    57457    59845: contig of 2389 bp in length
            *    59846 59945: gap of      100 bp
            *    59946    63972: contig of 4027 bp in length
            *    63973 64072: gap of      100 bp
            *    64073    82711: contig of 18639 bp in length
            *    82712 82811: gap of      100 bp
            *    82812   111814: contig of 29003 bp in length
            *   111815 111914: gap of      100 bp
            *   111915   120276: contig of 8362 bp in length
            *   120277 120376: gap of      100 bp
            *   120377   136660: contig of 16284 bp in length
            *   136661 136760: gap of      100 bp
            *   136761   153913: contig of 17153 bp in length
            *   153914 154013: gap of      100 bp
            *   154014   158659: contig of 4646 bp in length
            *   158660 158759: gap of      100 bp
            *   158760   164235: contig of 5476 bp in length
            *   164236 164335: gap of      100 bp
            *   164336   184996: contig of 20661 bp in length
            *   184997 185096: gap of      100 bp
            *   185097   193752: contig of 8656 bp in length.
FEATURES             Location/Qualifiers
     source          1. .193752
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-557F22"
                     /clone_lib="RPCI-11.2"
     misc_feature    1. .3250
                     /note="assembly_fragment:01752
                     fragment_chain:1"
     misc_feature    3351. .14600
                     /note="assembly_fragment:01177
                     fragment_chain:1"
     misc_feature    14701. .32357
                     /note="assembly_fragment:00673
                     fragment_chain:1"
     misc_feature    32458. .34886
                     /note="assembly_fragment:00884
                     fragment_chain:1"
     misc_feature    34987. .43490
                     /note="assembly_fragment:00652
                     fragment_chain:1"
     misc_feature    43591. .47437
                     /note="assembly_fragment:00628
                     fragment_chain:1"
     misc_feature    47538. .57356
                     /note="assembly_fragment:00983
                     fragment_chain:2"
     misc_feature    57457. .59845
                     /note="assembly_fragment:01129
                     fragment_chain:2"
     misc_feature    59946. .63972
                     /note="assembly_fragment:01207
                     fragment_chain:3"
     misc_feature    64073. .82711
                     /note="assembly_fragment:00301
                     fragment_chain:3"
     misc_feature    82812. .111814
                     /note="assembly_fragment:00470"
     misc_feature    111915. .120276
                     /note="assembly_fragment:00783.0"
     misc_feature    120377. .136660
                     /note="assembly_fragment:01014"
     misc_feature    136761. .153913
                     /note="assembly_fragment:01161"
     misc_feature    154014. .158659
                     /note="assembly_fragment:01329"
     misc_feature    158760. .164235
                     /note="assembly_fragment:01675"
     misc_feature    164336. .184996
                     /note="assembly_fragment:01884"
     misc_feature    185097. .193752
                     /note="assembly_fragment:01893"
BASE COUNT    55929 a  40409 c  39996 g  55686 t   1732 others
ORIGIN

  Query Match             100.0%;  Score 21;  DB 2;  Length 193752;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db  12585 GCTCTGACAACCTCAGGAAGG 12605


RESULT 12
AR036572/c
LOCUS       AR036572              246240 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 20 from patent US 5872237.
ACCESSION   AR036572
VERSION     AR036572.1  GI:5953240
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246240)
  AUTHORS   Feder,J.Nathan., Kronmal,G.Scott., Lauer,P.M., Ruddy,D.A.,
            Thomas,W., Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Megabase transcript map: novel sequences and antibodies thereto
  JOURNAL   Patent: US 5872237-A 20 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .246240
                     /organism="unknown"
BASE COUNT    73211 a  50177 c  50599 g  72252 t      1 others
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db 196370 GCTCTGACAACCTCAGGAAGG 196350


RESULT 13
AR036573/c
LOCUS       AR036573              246240 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 21 from patent US 5872237.
ACCESSION   AR036573
VERSION     AR036573.1  GI:5953241
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246240)
  AUTHORS   Feder,J.Nathan., Kronmal,G.Scott., Lauer,P.M., Ruddy,D.A.,
            Thomas,W., Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Megabase transcript map: novel sequences and antibodies thereto
  JOURNAL   Patent: US 5872237-A 21 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .246240
```

```
                     /organism="unknown"
BASE COUNT    73211 a  50177 c  50599 g  72252 t      1 others
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db 196370 GCTCTGACAACCTCAGGAAGG 196350


RESULT 14
AR036574/c
LOCUS       AR036574               246240 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 22 from patent US 5872237.
ACCESSION   AR036574
VERSION     AR036574.1  GI:5953242
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 246240)
  AUTHORS   Feder,J.Nathan., Kronmal,G.Scott., Lauer,P.M., Ruddy,D.A.,
            Thomas,W., Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Megabase transcript map: novel sequences and antibodies thereto
  JOURNAL   Patent: US 5872237-A 22 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .246240
                     /organism="unknown"
BASE COUNT    73211 a  50177 c  50599 g  72252 t      1 others
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 246240;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db 196370 GCTCTGACAACCTCAGGAAGG 196350


RESULT 15
HSU91328
LOCUS       HSU91328               246282 bp    DNA     linear   PRI 17-DEC-2001
DEFINITION  Human hereditary haemochromatosis region, histone 2A-like protein
            gene, hereditary haemochromatosis (HLA-H) gene, RoRet gene, and
            sodium phosphate transporter (NPT3) gene, complete cds.
ACCESSION   U91328
VERSION     U91328.1  GI:2088550
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 246282)
  AUTHORS   Ruddy,D.A., Kronmal,G.S., Lee,V.K., Mintier,G.A., Quintana,L.,
            Domingo,R. Jr., Meyer,N.C., Irrinki,A., McClelland,E.E., Fullan,A.,
            Mapa,F.A., Moore,T., Thomas,W., Loeb,D.B., Harmon,C.,
            Tsuchihashi,Z., Wolff,R.K., Schatzman,R.C. and Feder,J.N.
  TITLE     A 1.1-Mb transcript map of the hereditary hemochromatosis locus
  JOURNAL   Genome Res. 7 (5), 441-456 (1997)
  MEDLINE   97294057
   PUBMED   9149941
REFERENCE   2  (bases 1 to 246282)
  AUTHORS   Ruddy,D.A., Kronmal,G.S., Lee,V.K., Mintier,G.A., Quintana,L.,
            Domingo,R. Jr., Meyer,N.C., Irrinke,A., McClelland,E., Fullan,A.,
            Mapa,F.A., Moore,T., Thomas,W., Loeb,D.B., Harmon,C.,
            Tsuchihashi,Z., Wolff,R.K., Schatzman,R.C. and Feder,J.N.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-1997) Sequencing, Mercator Genetics, 4040
            Campbell Avenue, Menlo Park, CA 94025, USA
```

```
FEATURES             Location/Qualifiers
     source          1. .246282
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="pac222k22; pac75I14; bac132a2"
     repeat_region   complement(1231. .1501)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   3528. .3676
                     /rpt_family="MER3"
                     /rpt_type=dispersed
     repeat_region   4056. .4304
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   complement(5496. .5780)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   6032. .6461
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   8203. .8463
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   8507. .8616
                     /rpt_family="MSTAR"
                     /rpt_type=dispersed
     repeat_region   complement(8888. .9194)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   complement(9523. .9792)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   10109. .10375
                     /rpt_family="MSTAR"
                     /rpt_type=dispersed
     repeat_region   complement(10384. .10680)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   11482. .11742
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   complement(12899. .13326)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   15021. .15309
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   complement(15450. .15714)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     CDS             complement(16434. .16826)
                     /codon_start=1
                     /product="histone 2A-like protein"
                     /protein_id="AAB82086.1"
                     /db_xref="GI:2088554"
                     /translation="MSGRGKQGGKARAKAKSRSSRAGLQFPVGRVHRLLRKGNYAERV
                     GAGAPVYLAAVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLLGRVT
                     IAQGGVLPNIQAVLLPKKTESHHKAKGK"
     repeat_region   complement(18988. .19282)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   complement(20597. .20869)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   complement(24343. .24624)
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   25654. .25805
                     /rpt_family="MIR"
                     /rpt_type=dispersed
     repeat_region   29842. .29938
                     /rpt_family="Alu"
                     /rpt_type=dispersed
```

```
repeat_region   complement(30592. .30869)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   31223. .31507
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(31819. .32011)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   33810. .34150
                /rpt_family="MER1"
                /rpt_type=dispersed
repeat_region   34235. .34314
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(35768. .36062)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(36267. .36548)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   39132. .39252
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(40045. .40645)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   40960. .41246
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   41581. .42174
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   42432. .42541
                /rpt_family="MIR"
                /rpt_type=dispersed
repeat_region   44636. .44911
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(45532. .45851)
                /rpt_family="Alu"
                /rpt_type=dispersed
gene            complement(46833. .53618)
                /gene="HLA-H"
                /note="HFE"
CDS             complement(join(46833. .46873,47827. .47940,48099. .48374,
                49470. .49745,49955. .50218,53543. .53618))
                /gene="HLA-H"
                /codon_start=1
                /product="hereditary hemochromatosis"
                /protein_id="AAB82083.1"
                /db_xref="GI:2088551"
                /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
                WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTS
                SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGE
                EQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQG
                SRGAMGHYVLAERE"
repeat_region   complement(48868. .49182)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(50599. .50888)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(51145. .51372)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   51820. .51934
                /rpt_family="MER20"
                /rpt_type=dispersed
repeat_region   55276. .55719
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(56085. .56461)
                /rpt_family="L1"
                /rpt_type=dispersed
repeat_region   complement(56521. .56801)
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   57660. .57815
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   58118. .58483
                /rpt_family="MLT1"
                /rpt_type=dispersed
repeat_region   58672. .58900
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   complement(59299. .59737)
                /rpt_family="LTR2"
                /rpt_type=dispersed
repeat_region   complement(65665. .66013)
                /rpt_family="LTR2"
                /rpt_type=dispersed
repeat_region   66764. .67053
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region   67743. .68016
```

Query Match 100.0%; Score 21; DB 9; Length 246282;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db  49913 GCTCTGACAACCTCAGGAAGG 49933
```

Search completed: November 2, 2002, 05:39:51
Job time : 146.692 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:05:23 ; Search time 4.27747 Seconds
(without alignments)
8429.091 Million cell updates/sec

Title: US-09-981-606-4
Perfect score: 21
Sequence: 1 gctctgacaacctcaggaagg 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		21	100.0	21	21	AAA96771	PCR primer for his
c 2		21	100.0	5749	22	AAL36747	Human musculoskele
c 3		21	100.0	10825	18	AAT96690	Hereditary haemoch
c 4		21	100.0	10825	22	AAC68425	Human hereditary h
c 5		21	100.0	10825	22	AAC68426	Human hereditary h
c 6		21	100.0	10825	22	AAC68427	Human hereditary h
c 7		21	100.0	10825	22	AAC68428	Human hereditary h
c 8		21	100.0	12146	21	AAA96794	Genomic DNA of a h
9		21	100.0	235033	19	AAV57926	Hereditary haemoch
10		21	100.0	237326	19	AAV57903	Hereditary haemoch
c 11		16.8	80.0	409	22	AAI92117	Human polynucleoti
12		16.8	80.0	1743	21	AAC79809	Human secreted pro
c 13		16.8	80.0	1889	24	AAS94903	Human DNA sequence
c 14		16.8	80.0	1901	22	AAH02910	Human shear stress
c 15		16.8	80.0	1988	22	AAI93508	Human polynucleoti
16		16.8	80.0	4692	19	AAV04633	Porcine phosphoino
17		16.8	80.0	4692	20	AAV74099	Porcine G-protein
18		16.8	80.0	4692	21	AAZ86812	Pig p101 regulator
19		16.4	78.1	524	21	AAC55639	Human differential
20		16.4	78.1	1053	19	AAV34234	Human secreted pro
21		16.4	78.1	1056	21	AAZ98238	Human signal pepti
22		16.4	78.1	1067	22	AAI57811	Human polynucleoti
23		16.4	78.1	1075	19	AAV34235	Human secreted pro
24		16.4	78.1	1105	19	AAV34169	Human secreted pro
25		16.4	78.1	1145	22	AAI59597	Human polynucleoti
26		16.4	78.1	2081	22	AAS29823	Human cytoskeletal
c 27		16.4	78.1	2081	22	AAS29825	Human cytoskeletal
c 28		16.4	78.1	2083	22	AAS29824	Human cytoskeletal
29		16.2	77.1	269	22	AAK53906	Murine transcripti
30		16.2	77.1	700	22	AAH92065	Human inflammatory
31		16.2	77.1	700	22	AAH92066	Human inflammatory
32		16.2	77.1	1424	22	AAH21000	Bovine-derived DNA
33		16.2	77.1	1846	22	AAF88384	M. crenulata KLH2
34		16.2	77.1	1870	19	AAV60011	Nucleic acid A25co
35		16.2	77.1	2880	14	AAQ53218	Genomic sequence e
36		16.2	77.1	32189	22	AAS30115	Human lung antigen
37		16.2	77.1	32221	22	AAS30113	Human lung antigen
c 38		16	76.2	1174	21	AAC46275	Arabidopsis thalia
c 39		16	76.2	1317	21	AAC52106	Arabidopsis thalia
c 40		15.8	75.2	399	21	AAC02114	Human secreted pro
41		15.8	75.2	432	22	AAH98558	Human EST-derived
c 42		15.8	75.2	516	21	AAC75409	Human ORFX ORF964
c 43		15.8	75.2	650	19	AAV60069	Nucleic acid seque
c 44		15.8	75.2	690	18	AAT85180	Equine rhinovirus
45		15.8	75.2	700	22	AAH92019	Human inflammatory

ALIGNMENTS

RESULT 1
AAA96771
ID AAA96771 standard; DNA; 21 BP.
XX
AC AAA96771;
XX
DT 19-FEB-2001 (first entry)
XX
DE PCR primer for histocompatibility iron loading (HFE) gene exon 2.
XX
KW Human; histocompatibility iron loading protein; HFE protein;
KW major histocompatibility complex; non-classical class I gene;
KW chromosome 6p; iron disorder; haemochromatosis; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200058515-A1.
XX
PD 05-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US07982.
XX
PR 26-MAR-1999; 99US-0277457.
XX
PA (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI Rothenberg BE, Sawada-Hirai R, Barton JC;
XX
DR WPI; 2000-647244/62.
XX
PT Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT susceptibility to develop it, by determining the presence of a mutation

PT in exon 2 or an intron of a histocompatibility iron loading nucleic
PT acid -
XX
PS Claim 23; Page 5; 55pp; English.
XX
CC PCR primers A96770-71 were used to amplify a fragment of the human
CC histocompatibility iron loading (HFE) gene. The HFE gene is a major
CC histocompatibility (MHC) non-classical class I gene located on
CC chromosome 6p. Mutations in the gene lead to iron disorders. The
CC specification describes a method for diagnosing an iron disorder or a
CC genetic susceptibility to develop the disorder in a mammal. The method
CC comprises determining the presence of a mutation in exon 2 or an intron
CC of a HFE gene or protein. The mutation is not a C to G missense mutation
CC at nucleotide 187 of the sequence given in A96769 (Genbank Accession
CC number U60319). The presence of the mutation indicates the disorder or
CC the genetic susceptibility to the disorder. The method is used to
CC diagnose an iron disorder e.g. haemochromatosis, or a genetic
CC susceptibility to develop it.
XX
SQ Sequence 21 BP; 6 A; 6 C; 6 G; 3 T; 0 other;

Query Match 100.0%; Score 21; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCTGACAACCTCAGGAAGG 21
||||||||||||||||||||
Db 1 GCTCTGACAACCTCAGGAAGG 21

RESULT 2
AAL36747/c
ID AAL36747 standard; DNA; 5749 BP.
XX
AC AAL36747;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3112.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.

```
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-451937/48.
XX
PT   Isolated polypeptide for treating, preventing and/ or prognosing
PT   disorders related to the musculoskeletal system including
PT   musculoskeletal cancers and also for testing and detection e.g.
PT   diagnosis -
XX
PS   Example 2; SEQ ID NO 3112; 781pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC   (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC   for preventing, treating or ameliorating medical conditions e.g. by
CC   protein or gene therapy. The genes are isolated from a range of human
CC   tissues disclosed in the specification. The nucleic acids, proteins,
CC   antibodies and (ant)agonists are useful in the diagnosis, treatment
CC   and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC   other cancers of the adrenal gland, bone, bone marrow, breast,
CC   gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC   (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC   healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC   and (f) infectious diseases such as viral, bacterial, fungal and
CC   parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5749 BP; 1600 A; 1192 C; 1403 G; 1553 T; 1 other;

  Query Match             100.0%;  Score 21;  DB 22;  Length 5749;
  Best Local Similarity   100.0%;  Pred. No. 2.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           |||||||||||||||||||||
Db      67 GCTCTGACAACCTCAGGAAGG 47


RESULT 3
AAT96690/c
ID   AAT96690 standard; DNA; 10825 BP.
XX
AC   AAT96690;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis;
KW   gene therapy; prenatal screening; human; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             361..7147
FT                   /*tag=  a
FT                   /note= "contains introns"
FT   intron          437..3761
FT                   /*tag=  b
FT                   /number= 1
FT   intron          4026..4234
FT                   /*tag=  c
FT                   /number= 2
FT   intron          4511..5605
FT                   /*tag=  d
FT                   /number= 3
FT   intron          5882..6039
FT                   /*tag=  e
FT                   /number= 4
FT   intron          6154..7106
FT                   /*tag=  f
FT                   /number= 5
FT   mutation        3872
FT                   /*tag=  g
FT                   /note= "C to G substitution (24d2 mutation)
FT                          results in His to Asp substitution"
FT   variation       3878
FT                   /*tag=  h
FT                   /note= "A to T substitution (24d7 variant)
FT                          results in Ser to Cys substitution"
FT   mutation        5834
FT                   /*tag=  i
FT                   /note= "G to A substitution (24d1 mutation
FT                          associated with HH), results in Cys to
FT                          Tyr substitution"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US06254.
XX
PR   23-MAY-1996;   96US-0652265.
PR   04-APR-1996;   96US-0630912.
```

PR 16-APR-1996; 96US-0632673.
XX
PA (MERC-) MERCATOR GENETICS INC.
XX
PI Drayna DT, Feder JN, Gnirke A, Ruddy D, Thomas WJ;
PI Tsuchihashi Z, Wolff RK;
XX
DR WPI; 1997-512743/47.
DR P-PSDB; AAW36499.
XX
PT Hereditary haemochromatosis gene and variants - useful for diagnosis
PT and treatment of hereditary haemochromatosis disease
XX
PS Disclosure; Fig 3; 115pp; English.
XX
CC This genomic DNA sequence corresponds to the human gene whose
CC mutated form is associated with hereditary haemochromatosis (HH).
CC To identify this novel gene, allelic association patterns were
CC determined between known markers and the HH locus in the HLA region
CC of chromosome 6. A physical clone coverage was then generated
CC extending from D6S265, which is a marker that is centromeric of
CC HLA-A, in a telomeric direction through D6S276, a marker at which
CC the allelic association was no longer observed. A single mutation
CC (24d1) in the HH gene appears responsible for the majority of HH
CC disease. This comprises a G to A substitution that is present in
CC 86% of affected chromosomes and in 4% of unaffected chromosomes.
CC It results in a Cys to Tyr substitution in the encoded protein (see
CC AAW36499) at a critical disulphide bridge important for secondary
CC structure. The following are claimed: the HH genomic DNA (I), a
CC 1437 bp cDNA sequence (Ia) (see AAT96691) and their 24d1, 24d2 and
CC 24d7 variants; a cloning or expression vector; host cells; a
CC peptide product chosen from the HH gene product, its variants
CC (24d1, 24d2 and 24d7), or a peptide of at least 56 amino acid
CC residues of these; an antibody produced using the peptide; a method
CC to determine the presence or absence of the common HH gene
CC mutation; an animal model for the HH disease; metal chelation
CC agents, T-cell differentiation factors and therapeutic agents for
CC the mitigation of injury due to oxidative process in vivo or
CC mitigation of iron overload; a method for screening potential
CC therapeutic agents for activity in connection with HH disease; an
CC antisense oligonucleotide directed against a transcriptional
CC product of a nucleic acid sequence as above; and oligonucleotides
CC or pairs of oligonucleotides covering a range of nucleotides from
CC (I), (Ia) or their variants, useful for detecting a polymorphism in
CC the HH gene. The invention also relates to methods for screening
CC for HH homozygotes, to HH diagnosis, prenatal screening and
CC diagnosis, and therapies of HH disease, including gene therapy,
CC protein- and antibody-based therapeutics, and small molecule
CC therapeutics.
XX
SQ Sequence 10825 BP; 2996 A; 2254 C; 2648 G; 2927 T; 0 other;

Query Match 100.0%; Score 21; DB 18; Length 10825;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCTGACAACCTCAGGAAGG 21
||||||||||||||||||||
Db 4067 GCTCTGACAACCTCAGGAAGG 4047

RESULT 4
AAC68425/c
ID AAC68425 standard; DNA; 10825 BP.
XX
AC AAC68425;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36869.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2998 A; 2253 C; 2648 G; 2926 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCTGACAACCTCAGGAAGG 21
||||||||||||||||||||
Db 4067 GCTCTGACAACCTCAGGAAGG 4047

RESULT 5
AAC68426/c
ID AAC68426 standard; DNA; 10825 BP.
XX
AC AAC68426;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1 mutation DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;

PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36870.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2999 A; 2253 C; 2647 G; 2926 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCTGACAACCTCAGGAAGG 21
||||||||||||||||||||
Db 4067 GCTCTGACAACCTCAGGAAGG 4047

RESULT 6
AAC68427/c
ID AAC68427 standard; DNA; 10825 BP.
XX
AC AAC68427;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d2 mutation DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36871.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2998 A; 2252 C; 2649 G; 2926 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCTGACAACCTCAGGAAGG 21
||||||||||||||||||||
Db 4067 GCTCTGACAACCTCAGGAAGG 4047

RESULT 7
AAC68428/c
ID AAC68428 standard; DNA; 10825 BP.
XX
AC AAC68428;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1/2 mutation DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36872.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2999 A; 2252 C; 2648 G; 2926 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCTGACAACCTCAGGAAGG 21
||||||||||||||||||||
Db 4067 GCTCTGACAACCTCAGGAAGG 4047

RESULT 8
AAA96794/c
ID AAA96794 standard; cDNA; 12146 BP.

```
XX
AC   AAA96794;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Genomic DNA of a histocompatibility iron loading (HFE) gene.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            1028..1324
FT                   /*tag= a
FT                   /number= 1
FT   intron          1325..4651
FT                   /*tag= b
FT                   /number= 1
FT   exon            4652..4915
FT                   /*tag= c
FT                   /number= 2
FT   intron          4916..5124
FT                   /*tag= d
FT                   /number= 2
FT   exon            5125..5400
FT                   /*tag= e
FT                   /number= 3
FT   intron          5401..6493
FT                   /*tag= f
FT                   /number= 3
FT   exon            6494..6769
FT                   /*tag= g
FT                   /number= 4
FT   intron          6770..6927
FT                   /*tag= h
FT                   /number= 4
FT   exon            6928..7041
FT                   /*tag= i
FT                   /number= 5
FT   intron          7042..7994
FT                   /*tag= j
FT                   /number= 5
FT   exon            7995..9050
FT                   /*tag= k
FT                   /number= 6
FT   intron          9051..10205
FT                   /*tag= l
FT                   /number= 6
FT   exon            10206..10637
FT                   /*tag= m
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Example 1; Page 21-28; 55pp; English.
XX
CC   The present sequence represents the human histocompatibility iron
CC   loading (HFE) gene. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence 12146 BP; 3383 A; 2474 C; 2911 G; 3378 T; 0 other;

  Query Match             100.0%;  Score 21;  DB 21;  Length 12146;
  Best Local Similarity   100.0%;  Pred. No. 2.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           |||||||||||||||||||||
Db    4957 GCTCTGACAACCTCAGGAAGG 4937


RESULT 9
AAV57926
ID   AAV57926 standard; DNA; 235033 BP.
XX
AC   AAV57926;
XX
DT   23-DEC-1998  (first entry)
XX
DE   Hereditary haemochromatosis subregion from an unaffected individual.
XX
KW   Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
KW   diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;
KW   BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
KW   type 1 sodium transport gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO9814466-A1.
XX
PD   09-APR-1998.
XX
PF   30-SEP-1997;   97WO-US17658.
XX
PR   07-MAY-1997;   97US-0852495.
PR   01-OCT-1996;   96US-0724394.
XX
PA   (PROG-) PROGENTIOR INC.
XX
PI   Feder JN,  Kronmal GS,  Lauer PM,  Ruddy DA,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1998-240014/21.
XX
PT   Hereditary haemochromatosis gene products - used to develop products
PT   for the diagnosis and treatment of hereditary disorders in iron
PT   metabolism
XX
PS   Example 2; Fig 8; 209pp; English.
XX
CC   The present invention describes hereditary haemochromatosis gene
CC   products from the human haemochromatosis gene. The present sequence
CC   represents a hereditary haemochromatosis subregion from an individual
CC   unaffected by hereditary haemochromatosis (HH). Also described is a
CC   method to determine the presence or absence of the common hereditary
CC   haemochromatosis (HFE) gene mutation in an individual comprising:
CC   (a) providing DNA or RNA from the individual; and (b) assessing the
CC   DNA or RNA for the presence or absence of a haplotype or genotype where
CC   the presence or absence of the haplotype genotype indicates the likely
```

CC presence of the HFE gene mutation in the genome of the individual. The
CC HFE gene sequences from the present invention can be used to develop
CC products for use in the diagnosis and treatment of HFE. The present
CC invention also describes BTF genes, which are homologues of the milk
CC protein butyrophilin (BT), and can be used in the production of agonists
CC and antagonists of BT function. Also described are: (1) a RoRet gene
CC which can be used to develop products for the study, diagnosis and
CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
CC which are homologues of a type 1 sodium transport gene, and can
CC similarly be used for hypophosphatemia.
XX
SQ Sequence 235033 BP; 68786 A; 48466 C; 49441 G; 68340 T; 0 other;

Query Match 100.0%; Score 21; DB 19; Length 235033;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCTGACAACCTCAGGAAGG 21
||||||||||||||||||||
Db 43083 GCTCTGACAACCTCAGGAAGG 43103

RESULT 10
AAV57903
ID AAV57903 standard; DNA; 237326 BP.
XX
AC AAV57903;
XX
DT 21-DEC-1998 (first entry)
XX
DE Hereditary haemochromatosis subregion from an HH affected individual.
XX
KW Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
KW diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;
KW BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
KW type 1 sodium transport gene; ss.
XX
OS Homo sapiens.
XX
PN WO9814466-A1.
XX
PD 09-APR-1998.
XX
PF 30-SEP-1997; 97WO-US17658.
XX
PR 07-MAY-1997; 97US-0852495.
PR 01-OCT-1996; 96US-0724394.
XX
PA (PROG-) PROGENTIOR INC.
XX
PI Feder JN, Kronmal GS, Lauer PM, Ruddy DA, Thomas WJ;
PI Tsuchihashi Z, Wolff RK;
XX
DR WPI; 1998-240014/21.
XX
PT Hereditary haemochromatosis gene products - used to develop products
PT for the diagnosis and treatment of hereditary disorders in iron
PT metabolism
XX
PS Claim 1; Fig 9; 209pp; English.
XX
CC The present invention describes hereditary haemochromatosis gene
CC products from the human haemochromatosis gene. The present sequence
CC represents a hereditary haemochromatosis subregion from an hereditary
CC haemochromatosis (HH) affected individual. Also described is a
CC method to determine the presence or absence of the common hereditary
CC haemochromatosis (HFE) gene mutation in an individual comprising:
CC (a) providing DNA or RNA from the individual; and (b) assessing the
CC DNA or RNA for the presence or absence of a haplotype or genotype where
CC the presence or absence of the haplotype genotype indicates the likely
CC presence of the HFE gene mutation in the genome of the individual. The
CC HFE gene sequences from the present invention can be used to develop
CC products for use in the diagnosis and treatment of HFE. The present
CC invention also describes BTF genes, which are homologues of the milk
CC protein butyrophilin (BT), and can be used in the production of agonists
CC and antagonists of BT function. Also described are: (1) a RoRet gene
CC which can be used to develop products for the study, diagnosis and
CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
CC which are homologues of a type 1 sodium transport gene, and can
CC similarly be used for hypophosphatemia.
XX
SQ Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;

Query Match 100.0%; Score 21; DB 19; Length 237326;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCTGACAACCTCAGGAAGG 21
||||||||||||||||||||
Db 43033 GCTCTGACAACCTCAGGAAGG 43053

RESULT 11
AAI92117/c
ID AAI92117 standard; cDNA; 409 BP.
XX
AC AAI92117;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 12177.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR P-PSDB; AAO12186.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 1; SEQ ID NO 12177; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

```
SQ   Sequence 409 BP; 115 A; 99 C; 105 G; 87 T; 3 other;

  Query Match             80.0%;  Score 16.8;  DB 22;  Length 409;
  Best Local Similarity   90.0%;  Pred. No. 1.6e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 CTCTGACAACCTCAGGAAGG 21
           |||||| | ||||||||||||
Db     141 CTCTGAAATCCTCAGGAAGG 122


RESULT 12
AAC79809
ID   AAC79809 standard; cDNA; 1743 BP.
XX
AC   AAC79809;
XX
DT   12-FEB-2001  (first entry)
XX
DE   Human secreted protein gene 11 SEQ ID NO:21.
XX
KW   Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW   fungicide; ophthalmological; gene therapy; autoimmune disease; infection;
KW   hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW   cerebrovascular disorder; nervous system disorder; ocular disorder;
KW   wound healing; skin aging; food additive; preservative; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058336-A1.
XX
PD   05-OCT-2000.
XX
PF   23-MAR-2000; 2000WO-US07726.
XX
PR   26-MAR-1999;   99US-0126597.
PR   07-JAN-2000; 2000US-0174877.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Komatsoulis G;
XX
DR   WPI; 2000-602355/57.
DR   P-PSDB; AAB44772.
XX
PT   Nucleic acid encoding human secreted proteins, used to treat, prevent,
PT   ameliorate or diagnose medical conditions such as cancer, and
PT   autoimmune diseases -
XX
PS   Claim 1; Page 331-332; 391pp; English.
XX
CC   The polynucleotide sequences given in AAC79799 to AAC79848 encode the
CC   human secreted proteins given in AAB44762 to AAB44811. AAB44812 to
CC   AAB44829 represent human secreted polypeptide sequences and proteins
CC   homologous to them, which are used in the exemplification of the present
CC   invention. Human secreted proteins have activities based on the tissues
CC   and cells the genes are expressed in. Examples of activities are:
CC   immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
CC   cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
CC   neuroprotective; antibacterial; virucide; fungicide; and
CC   ophthalmological. The polynucleotides and polypeptides can be used to
CC   prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC   rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC   in diagnosing a pathological condition or susceptibility to a
CC   pathological condition. Disorders which are diagnosed or treated include
CC   autoimmune diseases, hyperproliferative disorders, cardiovascular
CC   disorders, cerebrovascular disorders, angiogenesis, nervous system
CC   disorders, infections caused by bacteria, viruses and fungi and ocular
CC   disorders. The polypeptides can also be used to aid wound healing and
CC   epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC   maintain organs before transplantation, for supporting cell culture of
CC   primary tissues, to regenerate tissues and in chemotaxis. The
CC   polypeptides can also be used as a food additive or preservative to
CC   increase or decrease storage capabilities. AAC79790 to AAC79798 and
CC   AAB44761 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence 1743 BP; 563 A; 396 C; 352 G; 423 T; 9 other;

  Query Match             80.0%;  Score 16.8;  DB 21;  Length 1743;
  Best Local Similarity   90.0%;  Pred. No. 1.8e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 CTCTGACAACCTCAGGAAGG 21
           |||||||||| ||||| |||
Db    1003 CTCTGACAACATCAGGTAGG 1022


RESULT 13
AAS94903/c
ID   AAS94903 standard; DNA; 1889 BP.
XX
AC   AAS94903;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Human DNA sequence #158 expressed during foam cell differentiation.
XX
KW   Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW   cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200177389-A2.
XX
PD   18-OCT-2001.
XX
PF   04-APR-2001; 2001WO-US11128.
XX
PR   05-APR-2000; 2000US-195106P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Shiffman D,  Somogyi R,  Lawn R,  Seilhamer JJ,  Porter GJ,  Mikita T;
PI   Tai J;
XX
DR   WPI; 2002-010925/01.
XX
PT   Composition useful for diagnosis of conditions, disorders or diseases
PT   associated with atherosclerosis, comprises several polynucleotides that
PT   are differentially expressed in foam cell development -
XX
PS   Claim 1; Page 213; 315pp; English.
XX
CC   The present invention relates to the isolation of human polynucleotide
CC   sequences that are differentially expressed during foam cell
CC   differentiation. The polynucleotide sequences of the invention or a
CC   composition comprising these polynucleotides are useful as a high
CC   throughput method for detecting altered expression of one or more
CC   polynucleotides in a sample. The polynucleotides can be used in the
CC   diagnosis of disorders associated with foam cell development such as
CC   atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC   coronary artery disease. The polynucleotide sequences can also be used
CC   as PCR primers and probes. The polynucleotides of the invention are also
CC   useful in gene therapy. AAS94746-AAS95021 represent the human
CC   polynucleotide sequences of the invention which are differentially
CC   expressed during foam cell differentiation.
XX
SQ   Sequence 1889 BP; 613 A; 354 C; 452 G; 442 T; 28 other;

  Query Match             80.0%;  Score 16.8;  DB 24;  Length 1889;
  Best Local Similarity   90.0%;  Pred. No. 1.8e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy       2 CTCTGACAACCTCAGGAAGG 21
           |||||| | ||||||||||||
Db     238 CTCTGAAATCCTCAGGAAGG 219


RESULT 14
AAH02910/c
ID   AAH02910 standard; DNA; 1901 BP.
XX
AC   AAH02910;
XX
DT   15-JUN-2001  (first entry)
XX
DE   Human shear stress-response coding sequence SEQ ID NO: 73.
XX
KW   Human; shear stress-response protein; vascular disease;
KW   arteriosclerosis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200125427-A1.
XX
PD   12-APR-2001.
XX
PF   02-OCT-2000; 2000WO-JP06840.
XX
PR   01-OCT-1999;   99JP-0280976.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
PA   (NOJI/) NOJIMA H.
XX
PI   Nojima H,  Yoshisue H,  Obayashi M,  Ota T,  Kawabata A,  Sakurada K;
PI   Kuga T,  Sekine S,  Nakamura Y,  Sugano S;
XX
DR   WPI; 2001-266308/27.
DR   P-PSDB; AAB90787.
XX
PT   DNA sequences, proteins encoded by them and antibodies against them
PT   useful in diagnosis and treatment of vascular disease caused by
PT   arteriosclerosis -
XX
PS   Claim 20; Page 422-425; 678pp; Japanese.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human shear stress response proteins. These are useful in the
CC   diagnosis, treatment and screening of vascular diseases caused by
CC   arteriosclerosis, including heart failure, post-PTCA restenosis and
CC   hypertension.
XX
SQ   Sequence 1901 BP; 592 A; 378 C; 460 G; 471 T; 0 other;

  Query Match             80.0%;  Score 16.8;  DB 22;  Length 1901;
  Best Local Similarity   90.0%;  Pred. No. 1.8e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 CTCTGACAACCTCAGGAAGG 21
           |||||| | ||||||||||||
Db     337 CTCTGAAATCCTCAGGAAGG 318


RESULT 15
AAI93508/c
ID   AAI93508 standard; cDNA; 1988 BP.
XX
AC   AAI93508;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 13568.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorders; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200164835-A2.
XX
PD   07-SEP-2001.
XX
PF   26-FEB-2001; 2001WO-US04927.
XX
PR   28-FEB-2000; 2000US-0515126.
PR   18-MAY-2000; 2000US-0577409.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-514838/56.
DR   P-PSDB; AAO13577.
XX
PT   Isolated nucleic acids and polypeptides, useful for preventing
PT   diagnosing and treating e.g. leukaemia, inflammation and immune
PT   disorders -
XX
PS   Claim 1; SEQ ID NO 13568; 1399pp + Sequence Listing; English.
XX
CC   The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC   the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1988 BP; 632 A; 384 C; 472 G; 500 T; 0 other;

  Query Match             80.0%;  Score 16.8;  DB 22;  Length 1988;
  Best Local Similarity   90.0%;  Pred. No. 1.8e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 CTCTGACAACCTCAGGAAGG 21
           |||||| | ||||||||||||
Db     337 CTCTGAAATCCTCAGGAAGG 318


Search completed: November  2, 2002, 04:19:57
Job time : 48.2775 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:55:33 ; Search time 0.832418 Seconds
(without alignments)
6196.774 Million cell updates/sec

Title: US-09-981-606-4
Perfect score: 21
Sequence: 1 gctctgacaacctcaggaagg 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	21	100.0	21	4	US-09-277-457-4	Sequence 4, Appli
c 2	21	100.0	10825	3	US-08-652-265-1	Sequence 1, Appli
c 3	21	100.0	10825	3	US-08-652-265-3	Sequence 3, Appli
c 4	21	100.0	10825	3	US-08-652-265-5	Sequence 5, Appli
c 5	21	100.0	10825	3	US-08-652-265-7	Sequence 7, Appli
c 6	21	100.0	10825	3	US-08-834-497A-1	Sequence 1, Appli
c 7	21	100.0	10825	3	US-08-834-497A-3	Sequence 3, Appli
c 8	21	100.0	10825	3	US-08-834-497A-5	Sequence 5, Appli
c 9	21	100.0	10825	3	US-08-834-497A-7	Sequence 7, Appli
c 10	21	100.0	10825	4	US-09-503-444A-1	Sequence 1, Appli
c 11	21	100.0	10825	4	US-09-503-444A-3	Sequence 3, Appli
c 12	21	100.0	10825	4	US-09-503-444A-5	Sequence 5, Appli
c 13	21	100.0	10825	4	US-09-503-444A-7	Sequence 7, Appli
c 14	21	100.0	12146	4	US-09-277-457-27	Sequence 27, Appl
c 15	21	100.0	246240	2	US-08-724-394A-20	Sequence 20, Appl
c 16	21	100.0	246240	2	US-08-724-394A-21	Sequence 21, Appl
c 17	21	100.0	246240	2	US-08-724-394A-22	Sequence 22, Appl
18	16.8	80.0	732	4	US-08-998-416-903	Sequence 903, App
19	16.8	80.0	4692	2	US-08-916-917-1	Sequence 1, Appli
20	16.8	80.0	4692	2	US-08-972-631-1	Sequence 1, Appli
21	16.8	80.0	4692	2	US-08-972-629-1	Sequence 1, Appli
22	16.8	80.0	4692	2	US-08-972-630-1	Sequence 1, Appli
23	16.8	80.0	4692	2	US-08-672-211-1	Sequence 1, Appli
24	16.8	80.0	4692	3	US-09-225-170-1	Sequence 1, Appli
25	16.2	77.1	1870	1	US-08-592-126-86	Sequence 86, Appl
26	16.2	77.1	2880	1	US-08-158-189-1	Sequence 1, Appli
c 27	15.8	75.2	650	1	US-08-592-126-151	Sequence 151, App
c 28	15.8	75.2	690	4	US-09-091-219-8	Sequence 8, Appli
c 29	15.8	75.2	2995	1	US-08-592-126-85	Sequence 85, Appl
c 30	15.8	75.2	7278	4	US-09-091-219-1	Sequence 1, Appli
c 31	15.2	72.4	597	4	US-08-991-789A-186	Sequence 186, App
c 32	15.2	72.4	597	4	US-09-062-451-186	Sequence 186, App
c 33	15.2	72.4	4035	6	5198359-1	Patent No. 5198359
c 34	15.2	72.4	4035	6	5449756-1	Patent No. 5449756
c 35	15.2	72.4	4582	2	US-08-993-228-9	Sequence 9, Appli
c 36	14.8	70.5	512	3	US-08-545-809A-2	Sequence 2, Appli
c 37	14.8	70.5	574	2	US-08-836-943-3	Sequence 3, Appli
38	14.8	70.5	702	1	US-07-938-333A-4	Sequence 4, Appli
39	14.8	70.5	702	1	US-08-660-216A-4	Sequence 4, Appli
c 40	14.8	70.5	704	4	US-08-896-164-49	Sequence 49, Appl
41	14.8	70.5	1147	1	US-08-417-103-15	Sequence 15, Appl
42	14.8	70.5	1244	1	US-07-816-283-7	Sequence 7, Appli
43	14.8	70.5	1244	1	US-08-417-103-7	Sequence 7, Appli
44	14.8	70.5	1351	1	US-07-816-283-5	Sequence 5, Appli
45	14.8	70.5	1351	1	US-08-417-103-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-277-457-4
; Sequence 4, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
; APPLICANT: Rothenberg, Barry E.
; APPLICANT: Sawada-Hirai, Ritsuko
; APPLICANT: Barton, James C.
; TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
; FILE REFERENCE: 10653/002001
; CURRENT APPLICATION NUMBER: US/09/277,457
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Reverse Primer
US-09-277-457-4

Query Match 100.0%; Score 21; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCTGACAACCTCAGGAAGG 21
||||||||||||||||||||
Db 1 GCTCTGACAACCTCAGGAAGG 21

RESULT 2
US-08-652-265-1/c
; Sequence 1, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California

```
;       COUNTRY:  USA
;       ZIP:  94111-3834
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/652,265
;       FILING DATE:  23-MAY-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Smith, William M.
;       REGISTRATION NUMBER:  30,223
;       REFERENCE/DOCKET NUMBER:  17957-000500
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 576-0200
;       TELEFAX:  (415) 576-0300
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10825 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;       LOCATION:  6040..6153, 7107..7147)
;       OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /note= "No. 6025130mal or wild-type (unaffected)
;       OTHER INFORMATION:  Hereditary Hemochromatosis (HH) gene
;       OTHER INFORMATION:  allele"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  140..7319
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) allele
;       OTHER INFORMATION:  cDNA (SEQ ID NO:9)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  3852..3891
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d2(C)
;       OTHER INFORMATION:  allele (SEQ ID NO:41)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  5507..6023
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d1(G)
;       OTHER INFORMATION:  allele (SEQ ID NO:20)"
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3872, "c")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3878, "a")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d7
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "g")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d1
US-08-652-265-1

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           |||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 3
US-08-652-265-3/c
; Sequence 3, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d1 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d1 allele cDNA (SEQ ID NO:10)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
```

```
;      OTHER INFORMATION:  for 24d2(C) allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-3

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db   4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 4
US-08-652-265-5/c
; Sequence 5, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d2 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d2 allele cDNA (SEQ ID NO:11)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(G) allele (SEQ ID NO:20)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-08-652-265-5

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db   4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 5
US-08-652-265-7/c
; Sequence 7, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  and 24d2 mutations"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  cDNA containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles
;      OTHER INFORMATION:  (SEQ ID NO:12)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-7

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           |||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 6
US-08-834-497A-1/c
; Sequence 1, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /note= "No. 6140305mal or wild-type (unaffected)
;      OTHER INFORMATION:  Hereditary Hemochromatosis (HH) gene
;      OTHER INFORMATION:  allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) allele
;      OTHER INFORMATION:  cDNA (SEQ ID NO:9)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;      OTHER INFORMATION:  sequence surrounding variant for 24d2(C)
;      OTHER INFORMATION:  allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;      OTHER INFORMATION:  sequence surrounding variant for 24d1(G)
```

```
;       OTHER INFORMATION:  allele (SEQ ID NO:20)"
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3872, "c")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3878, "a")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d7
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "g")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d1
US-08-834-497A-1

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db   4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 7
US-08-834-497A-3/c
; Sequence 3, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d1 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d1 allele cDNA (SEQ ID NO:10)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(C) allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-834-497A-3

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db   4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 8
US-08-834-497A-5/c
; Sequence 5, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
```

```
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d2 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d2 allele cDNA (SEQ ID NO:11)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(G) allele (SEQ ID NO:20)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
```

```
US-08-834-497A-5

  Query Match           100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity  100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           |||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 9
US-08-834-497A-7/c
; Sequence 7, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
```

```
;       OTHER INFORMATION:
;       OTHER INFORMATION:  and 24d2 mutations"
;       OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;       OTHER INFORMATION:  gene containing a combination of both
;       OTHER INFORMATION:  24d1 and 24d2 alleles"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  140..7319
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  cDNA containing a combination of both
;       OTHER INFORMATION:  24d1 and 24d2 alleles
;       OTHER INFORMATION:  (SEQ ID NO:12)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  3852..3891
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  genomic sequence surrounding variant
;       OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  5507..6023
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  genomic sequence surrounding variant
;       OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3872, "g")
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "a")
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d1
US-08-834-497A-7

  Query Match             100.0%;  Score 21;  DB 3;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           |||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 10
US-09-503-444A-1/c
; Sequence 1, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;     APPLICANT:  Thomas, Winston J.
;     APPLICANT:  Drayna, Dennis T.
;     APPLICANT:  Feder, John N.
;     APPLICANT:  Gnirke, Andreas
;     APPLICANT:  Ruddy, David
;     APPLICANT:  Tsuchihashi, Zenta
;     APPLICANT:  Wolff, Roger K.
;     TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;     NUMBER OF SEQUENCES:  44
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Pennie & Edmonds LLP
;       STREET:  1155 Avenue of the Americas
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows 95
;       SOFTWARE:  WordPerfect Version 8
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/503,444A
;       FILING DATE:  14-Feb-2000
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/652,265
;       FILING DATE:  23-May-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/632,673
;       FILING DATE:  16-Apr-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/630,912
;       FILING DATE:  04-Apr-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Poissant, Brian M.
;       REGISTRATION NUMBER:  28,462
;       REFERENCE/DOCKET NUMBER:  8907-0088-999
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-790-9090
;       TELEFAX:  212-869-9741
;       TELEX:  66141
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10825 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;       LOCATION:  6040..6153, 7107..7147)
;       OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /note= "No. 6228594mal or wild-type (unaffected)
;       OTHER INFORMATION:  Hereditary Hemochromatosis (HH) gene
;       OTHER INFORMATION:  allele"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  140..7319
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) allele
;       OTHER INFORMATION:  cDNA (SEQ ID NO:9)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  3852..3891
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d2(C)
;       OTHER INFORMATION:  allele (SEQ ID NO:41)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  5507..6023
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d1(G)
;       OTHER INFORMATION:  allele (SEQ ID NO:20)"
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3872, "c")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3878, "a")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d7
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "g")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
```

```
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-1

  Query Match             100.0%;  Score 21;  DB 4;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           |||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 11
US-09-503-444A-3/c
; Sequence 3, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d1 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d1 allele cDNA (SEQ ID NO:10)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(C) allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-3

  Query Match             100.0%;  Score 21;  DB 4;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           |||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 12
US-09-503-444A-5/c
; Sequence 5, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d2 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d2 allele cDNA (SEQ ID NO:11)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(G) allele (SEQ ID NO:20)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-09-503-444A-5

  Query Match             100.0%;  Score 21;  DB 4;  Length 10825;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           |||||||||||||||||||||
Db    4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 13
US-09-503-444A-7/c
; Sequence 7, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  and 24d2 mutations"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  cDNA containing a combination of both
;      OTHER INFORMATION:  24d1 and 24d2 alleles
;      OTHER INFORMATION:  (SEQ ID NO:12)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(G) allele (SEQ ID NO:42)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
```

```
;      LOCATION:  replace(3872, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-7

  Query Match            100.0%;  Score 21;  DB 4;  Length 10825;
  Best Local Similarity  100.0%;  Pred. No. 0.24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db   4067 GCTCTGACAACCTCAGGAAGG 4047


RESULT 14
US-09-277-457-27/c
; Sequence 27, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
;  APPLICANT: Rothenberg, Barry E.
;  APPLICANT:  Sawada-Hirai, Ritsuko
;  APPLICANT:  Barton, James C.
;  TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
;  FILE REFERENCE: 10653/002001
;  CURRENT APPLICATION NUMBER: US/09/277,457
;  CURRENT FILING DATE:  1999-03-26
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
;   LENGTH: 12146
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-09-277-457-27

  Query Match            100.0%;  Score 21;  DB 4;  Length 12146;
  Best Local Similarity  100.0%;  Pred. No. 0.25;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db   4957 GCTCTGACAACCTCAGGAAGG 4937


RESULT 15
US-08-724-394A-20/c
; Sequence 20, Application US/08724394A
; Patent No. 5872237
;  GENERAL INFORMATION:
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Kronmal, Gregory S.
;    APPLICANT:  Lauer, Peter M.
;    APPLICANT:  Ruddy, David A.
;    APPLICANT:  Thomas, Winston
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Megabase Transcript Map:  No. 5872237el
;    TITLE OF INVENTION:  Sequences and Antibodies Thereto
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,394A
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  017957-000100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  415-576-0300
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  246240 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..246240
;      OTHER INFORMATION:  /note= "HLA-H.CONTIG"
US-08-724-394A-20

  Query Match            100.0%;  Score 21;  DB 2;  Length 246240;
  Best Local Similarity  100.0%;  Pred. No. 0.37;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||||||
Db 196370 GCTCTGACAACCTCAGGAAGG 196350


Search completed: November  2, 2002, 06:45:00
Job time : 31.8324 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:10:18 ; Search time 31.772 Seconds
(without alignments)
8920.945 Million cell updates/sec

Title: US-09-981-606-4
Perfect score: 21
Sequence: 1 gctctgacaacctcaggaagg 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	19.4	92.4	1055	12	CNS05AO2	AL328763 Tetraodon
	2	17.8	84.8	286	9	AL507302	AL507302 AL507302
	3	17.8	84.8	384	10	BE403663	BE403663 WHE0435_C
	4	17.8	84.8	399	10	BE638057	BE638057 WHE0995-0
	5	17.8	84.8	436	10	BF857517	BF857517 RC5-FT019
	6	17.8	84.8	446	10	BM374748	BM374748 EBma05_SQ
	7	17.8	84.8	446	10	BM377929	BM377929 EBem04_SQ
	8	17.8	84.8	447	12	CNS03FKJ	AL241804 Tetraodon
	9	17.8	84.8	460	10	BF485164	BF485164 WHE1789_B
	10	17.8	84.8	462	10	BF474007	BF474007 WHE0839_H
	11	17.8	84.8	466	10	BE415673	BE415673 MWL037.H0
	12	17.8	84.8	467	9	AV913575	AV913575 AV913575
	13	17.8	84.8	470	10	BG263010	BG263010 WHE0939_G
	14	17.8	84.8	492	10	BF630175	BF630175 HVSMEb000
	15	17.8	84.8	499	10	BE406885	BE406885 WHE0433_d
	16	17.8	84.8	506	10	BF630173	BF630173 HVSMEb000
	17	17.8	84.8	508	10	BF630174	BF630174 HVSMEb000
	18	17.8	84.8	516	10	BE602911	BE602911 HVSMEh010
	19	17.8	84.8	517	10	BF483499	BF483499 WHE2334_F
	20	17.8	84.8	518	10	BE638022	BE638022 WHE0995-0
c	21	17.8	84.8	537	12	AZ583590	AZ583590 1M0378P02
	22	17.8	84.8	542	10	BF260641	BF260641 HVSMEf002
	23	17.8	84.8	543	10	BE494594	BE494594 WHE1255_E
	24	17.8	84.8	544	10	BM137873	BM137873 WHE0472_E
	25	17.8	84.8	553	10	BE494735	BE494735 WHE1258_H
	26	17.8	84.8	555	10	BE637786	BE637786 WHE1755-1
	27	17.8	84.8	561	10	BF484643	BF484643 WHE2318_A
	28	17.8	84.8	572	9	AV939721	AV939721 AV939721
	29	17.8	84.8	572	10	BM138377	BM138377 WHE0491_H
	30	17.8	84.8	573	10	BE499545	BE499545 WHE0961_C
	31	17.8	84.8	574	10	BE418029	BE418029 SCL011.C1
	32	17.8	84.8	580	10	BF484242	BF484242 WHE2309_H
	33	17.8	84.8	591	10	BF258579	BF258579 HVSMEf001
	34	17.8	84.8	604	10	BE419523	BE419523 WWS014.B2
	35	17.8	84.8	618	9	AV916007	AV916007 AV916007
	36	17.8	84.8	623	9	AV914415	AV914415 AV914415
	37	17.8	84.8	635	10	BI958641	BI958641 HVSMEn001
	38	17.8	84.8	640	10	BF473018	BF473018 WHE0927_C
	39	17.8	84.8	641	10	BG345151	BG345151 HVSMEg001
	40	17.8	84.8	642	10	BE601540	BE601540 HVSMEh009
	41	17.8	84.8	646	10	BE430692	BE430692 SUN007.B0
	42	17.8	84.8	649	10	BE430652	BE430652 SUN005.C1
	43	17.8	84.8	650	10	BI958621	BI958621 HVSMEn001
	44	17.8	84.8	660	10	BI960274	BI960274 HVSMEn002
	45	17.8	84.8	660	10	BE497139	BE497139 WHE0764_G

ALIGNMENTS

RESULT 1
CNS05AO2
LOCUS CNS05AO2 1055 bp DNA linear GSS 26-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence SP6 end of clone 001B13 of library B from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL328763
VERSION AL328763.1 GI:8222385
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1055)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Charaterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1055)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1055)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.
FEATURES Location/Qualifiers
source 1. .1055
/organism="Tetraodon nigroviridis"

```
                     /db_xref="taxon:99883"
                     /clone="001B13"
                     /clone_lib="B"
                     /note="Genoscope sequence ID : C0AB001CA07B1~end : SP6"
BASE COUNT       258 a    283 c    236 g    261 t     17 others
ORIGIN

  Query Match             92.4%;  Score 19.4;  DB 12;  Length 1055;
  Best Local Similarity   95.2%;  Pred. No. 1e+02;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||||||||| |||
Db    768 GCTCTGACAACCTCAGGCAGG 788


RESULT 2
AL507302
LOCUS       AL507302                286 bp    mRNA    linear   EST 04-JAN-2001
DEFINITION  AL507302 Hordeum vulgare Barke developing caryopsis (3.-15.DAP)
            Hordeum vulgare cDNA clone HY01M06V 5', mRNA sequence.
ACCESSION   AL507302
VERSION     AL507302.1  GI:12033517
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 286)
  AUTHORS   Michalek,W., Weschke,W., Pleissner,K.-P. and Graner,A.
  TITLE     EST sequencing and analysis in barley
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Michalek W
            Institute for Plant Genetics and Crop Plant Research
            Corrensstr.3, D-06466 Gatersleben, Germany
            Email: michalek@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de
            Seq primer: T3 primer for 5'end.
FEATURES             Location/Qualifiers
     source          1. .286
                     /organism="Hordeum vulgare"
                     /cultivar="Barke"
                     /db_xref="taxon:4513"
                     /clone="HY01M06V"
                     /clone_lib="Hordeum vulgare Barke developing caryopsis
                     (3.-15.DAP)"
                     /tissue_type="developing caryopsis (3.-15.DAP)"
                     /lab_host="XLOLR"
                     /note="Vector: plasmid pBK-CMV; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was made from developing caryopsis (3.-15.DAP)
                     of spring barley variety 'Barke', a high quality malting
                     variety. Cloning sites: EcoRI (5'-end of cDNA) and XhoI
                     (3'-end of cDNA). NOTE: Due to a cloning artefact caused
                     by the kit, in most cases the EcoRI site is NOT present,
                     as well as the EcoRI adapter. Average insert size is 1 kb
                     Sequence trimming: Vector sequences and sequence ends were
                     trimmed from the 5'-and 3'-end until a 50 bp window
                     contains less than two ambiguities. The maximum length was
                     set to 700 bp"
BASE COUNT       61 a     81 c     77 g     65 t      2 others
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 286;
  Best Local Similarity   90.5%;  Pred. No. 3.8e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||| |||||||| |||
Db     54 GCTCTGACCACCTCAGGCAGG 74


RESULT 3
BE403663
LOCUS       BE403663                384 bp    mRNA    linear   EST 21-JUL-2000
DEFINITION  WHE0435_C03_F05ZS Wheat etiolated seedling root cDNA library
            Triticum aestivum cDNA clone WHE0435_C03_F05, mRNA sequence.
ACCESSION   BE403663
VERSION     BE403663.1  GI:9363131
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 384)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Strategene SK primer.
FEATURES             Location/Qualifiers
     source          1. .384
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE0435_C03_F05"
                     /clone_lib="Wheat etiolated seedling root cDNA library"
                     /tissue_type="Root"
                     /dev_stage="Five day old etiolated seedling"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized
                     , germinated and grown aseptically in the  dark at room
                     temperature on filter paper with water, nystatin and
                     cefotaxime in covered crystallization dishes. Roots were
                     harvested.  The tissue, total RNA, and poly(A) RNA were
                     prepared, a cDNA library  was made, and the cDNA clones
                     were in vivo excised to give pBluescript  phagemids in the
                     TJ Close lab (Choi, Close, Fenton) at the University  of
                     California, Riverside.  Plasmid DNA preparations and DNA
                     sequencing  were performed in the OD Anderson lab (all
                     other authors)."
BASE COUNT       75 a    107 c    104 g     98 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 10;  Length 384;
  Best Local Similarity   90.5%;  Pred. No. 4.2e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||| |||||||| |||
Db     24 GCTCTGACCACCTCAGGCAGG 44


RESULT 4
BE638057
LOCUS       BE638057                399 bp    mRNA    linear   EST 25-AUG-2000
DEFINITION  WHE0995-0998_C17_C17ZS Wheat pre-anthesis spike cDNA library
            Triticum aestivum cDNA clone WHE0995-0998_C17_C17, mRNA sequence.
ACCESSION   BE638057
VERSION     BE638057.1  GI:9921168
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Pre-anthesis spike cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1. .399
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE0995-0998_C17_C17"
                     /clone_lib="Wheat pre-anthesis spike cDNA library"
                     /tissue_type="Spike before anthesis"
                     /dev_stage="Adult plant"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
                     greenhouse.  Whole spike with awns trimmed,  white, green
                     and yellow anther were collected and total RNA, and
                     poly(A) RNA were prepared, a cDNA library was made, and
                     the cDNA  clones were in vivo excised to give pBluescript
                     phagemids in the TJ  Close lab (Choi, Close, Fenton) at
                     the University of California,  Riverside.  Plasmid DNA
                     preparations and DNA sequencing were performed  in the OD
                     Anderson lab (all other authors)."
BASE COUNT       72 a    132 c    111 g     84 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 10;  Length 399;
  Best Local Similarity   90.5%;  Pred. No. 4.3e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          ||||||| |||||||| |||
Db    214 GCTCTGACCACCTCAGGCAGG 234


RESULT 5
BF857517
LOCUS       BF857517                 436 bp    mRNA    linear   EST 16-JAN-2001
DEFINITION  RC5-FT0195-031100-021-G04 FT0195 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF857517
VERSION     BF857517.1  GI:12245261
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 436)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-FT0195-
            031100-021-G04&t3=2000-11-03&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 11
            High quality sequence stop: 435.
FEATURES             Location/Qualifiers
     source          1. .436
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="FT0195"
                     /dev_stage="Adult"
                     /note="Organ: prostate_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      123 a     83 c     79 g    151 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 10;  Length 436;
  Best Local Similarity   90.5%;  Pred. No. 4.4e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          || |||||||||||||||| |
Db     70 GCACTGACAACCTCAGGAATG 90


RESULT 6
BM374748
LOCUS       BM374748                 446 bp    mRNA    linear   EST 10-JAN-2002
DEFINITION  EBma05_SQ002_L01_R IGF Barley EBma05 library Hordeum vulgare cDNA
            clone EBma05_SQ002_L01 5', mRNA sequence.
ACCESSION   BM374748
VERSION     BM374748.1  GI:18118138
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 446)
  AUTHORS   Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
            Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
  TITLE     Development of Barley Transcriptome Resources
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Waugh R
            Unit of Genomics
            Scottish Crop Research Institute
            Invergowrie, Dundee, DD2 5DA, Scotland, UK
            Tel: 00 44 1382 562731
            Fax: 00 44 1382 562426
            Email: rwaugh@scri.sari.ac.uk
            All sequence has a Phred quality score of 20 or over
            Seq primer: M13 reverse.
FEATURES             Location/Qualifiers
     source          1. .446
                     /organism="Hordeum vulgare"
                     /cultivar="Optic"
                     /db_xref="taxon:4513"
                     /clone="EBma05_SQ002_L01"
```

```
                     /clone_lib="IGF Barley EBma05 library"
                     /tissue_type="Maternal tissue"
                     /dev_stage="12 days post anthesis"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
                     Non-normalised library, directionally cloned into pSPORT1.
                     Derived from maternal tissue dissected from developing
                     grains (12 days post anthesis) in glasshouse grown barley
                     plants.  Developed as part of the barley transcriptome
                     resources of BBSRC/SEERAD funded cereal IGF (Investigating
                     Gene Function) project."
BASE COUNT       83 a    147 c    124 g     91 t      1 others
ORIGIN

  Query Match            84.8%;  Score 17.8;  DB 10;  Length 446;
  Best Local Similarity  90.5%;  Pred. No. 4.4e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||| |||||||| |||
Db    238 GCTCTGACCACCTCAGGCAGG 258


RESULT 7
BM377929
LOCUS       BM377929                446 bp    mRNA    linear   EST 10-JAN-2002
DEFINITION  EBem04_SQ004_K01_R IGF Barley EBem04 library Hordeum vulgare cDNA
            clone EBem04_SQ004_K01 5', mRNA sequence.
ACCESSION   BM377929
VERSION     BM377929.1  GI:18121319
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 446)
  AUTHORS   Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
            Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
  TITLE     Development of Barley Transcriptome Resources
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Waugh R
            Unit of Genomics
            Scottish Crop Research Institute
            Invergowrie, Dundee, DD2 5DA, Scotland, UK
            Tel: 00 44 1382 562731
            Fax: 00 44 1382 562426
            Email: rwaugh@scri.sari.ac.uk
            All sequence has a Phred quality score of 20 or over
            Seq primer: M13 reverse.
FEATURES             Location/Qualifiers
     source          1. .446
                     /organism="Hordeum vulgare"
                     /cultivar="Optic"
                     /db_xref="taxon:4513"
                     /clone="EBem04_SQ004_K01"
                     /clone_lib="IGF Barley EBem04 library"
                     /tissue_type="Embryo"
                     /dev_stage="12 days post anthesis"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
                     Non-normalised library, directionally cloned into pSPORT1.
                     Derived from embryos dissected from developing grains (12
                     days post anthesis) in glasshouse grown barley plants.
                     Developed as part of the barley transcriptome resources of
                     BBSRC/SEERAD funded cereal IGF (Investigating Gene
                     Function) project."
BASE COUNT       92 a    122 c    128 g    104 t
ORIGIN

  Query Match            84.8%;  Score 17.8;  DB 10;  Length 446;
  Best Local Similarity  90.5%;  Pred. No. 4.4e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||| |||||||| |||
Db     45 GCTCTGACCACCTCAGGCAGG 65


RESULT 8
CNS03FKJ
LOCUS       CNS03FKJ                447 bp    DNA     linear   GSS 17-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            022012 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL241804
VERSION     AL241804.1  GI:7962573
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 447)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 447)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="022012"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG022BH06LP1~end : T7"
BASE COUNT      133 a    126 c    114 g     67 t      7 others
ORIGIN

  Query Match            84.8%;  Score 17.8;  DB 12;  Length 447;
  Best Local Similarity  90.5%;  Pred. No. 4.4e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||||||| ||| |||
Db    244 GCTCTGACAACCTGAGGCAGG 264


RESULT 9
BF485164
LOCUS       BF485164                460 bp    mRNA    linear   EST 06-DEC-2000
DEFINITION  WHE1789_B04_D07ZS Wheat pre-anthesis spike cDNA library Triticum
            aestivum cDNA clone WHE1789_B04_D07, mRNA sequence.
ACCESSION   BF485164
VERSION     BF485164.1  GI:11568465
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Pre-anthesis spike cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE1789_B04_D07"
                     /clone_lib="Wheat pre-anthesis spike cDNA library"
                     /tissue_type="Spike before anthesis"
                     /dev_stage="Adult plant"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
                     greenhouse.  Whole spike with awns trimmed,  white, green
                     and yellow anther were collected and total RNA, and
                     poly(A) RNA were prepared, a cDNA library was made, and
                     the cDNA  clones were in vivo excised to give pBluescript
                     phagemids in the TJ  Close lab (Choi, Close, Fenton) at
                     the University of California,  Riverside.  Plasmid DNA
                     preparations and DNA sequencing were performed  in the OD
                     Anderson lab (all other authors)."
BASE COUNT       90 a    146 c    121 g    102 t      1 others
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 10;  Length 460;
  Best Local Similarity   90.5%;  Pred. No. 4.4e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           ||||||||  ||||||||  |||
Db     151 GCTCTGACCACCTCAGGCAGG 171


RESULT 10
BF474007
LOCUS       BF474007                462 bp    mRNA    linear   EST 04-DEC-2000
DEFINITION  WHE0839_H09_P17ZS Wheat vernalized crown cDNA library Triticum
            aestivum cDNA clone WHE0839_H09_P17, mRNA sequence.
ACCESSION   BF474007
VERSION     BF474007.1  GI:11543189
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Vernalized crown cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE0839_H09_P17"
                     /clone_lib="Wheat vernalized crown cDNA library"
                     /tissue_type="Crown tissue of seedling"
                     /dev_stage="Five-week old seedling"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Seeds were germinated and
                     grown at 4 C for 5 weeks.  The tissue, total  RNA, and
                     poly(A) RNA were prepared, a cDNA library was made, and
                     the  cDNA clones were in vivo excised to give pBluescript
                     phagemids in the  TJ Close lab (Choi, Close, Fenton) at
                     the University of California,  Riverside.  Plasmid DNA
                     preparations and DNA sequencing were performed  in the OD
                     Anderson lab (all other authors)."
BASE COUNT      104 a    118 c    129 g    111 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 10;  Length 462;
  Best Local Similarity   90.5%;  Pred. No. 4.4e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           ||||||||  ||||||||  |||
Db      26 GCTCTGACCACCTCAGGCAGG 46


RESULT 11
BE415673
LOCUS       BE415673                466 bp    mRNA    linear   EST 24-JUL-2000
DEFINITION  MWL037.H09000424 ITEC MWL Wheat Root Library Triticum aestivum cDNA
            clone MWL037.H09, mRNA sequence.
ACCESSION   BE415673
VERSION     BE415673.1  GI:9413519
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 466)
  AUTHORS   Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
            ,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
            Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
            Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
            Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
            Sorrells,M., Warburton,M. and Wenzel,G.
  TITLE     International Triticeae EST Cooperative (ITEC): Production of
            Expressed Sequence Tags for Species of the Triticeae
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Warburton M
            Applied Biotechnology Center, CIMMYT
            Apdo. Postal 6-641, 06600 Mexico DF MEXICO
            Tel: 52-5-7269091 ext 1381
            Fax: 52-5-7267558/59
            Email: mwarburton@cgnet.com
            International Triticeae EST Cooperative (ITEC)
            http://wheat.pw.usda.gov/genome.
FEATURES             Location/Qualifiers
     source          1. .466
                     /organism="Triticum aestivum"
```

```
                     /cultivar="Atlas"
                     /db_xref="taxon:4565"
                     /clone="MWL037.H09"
                     /clone_lib="ITEC MWL Wheat Root Library"
                     /tissue_type="root"
                     /dev_stage="8 day old"
                     /note="Vector: pYES2 (Invitrogen); 0.5-1.5 Kbp average
                     insert size."
BASE COUNT       81 a    164 c    136 g     82 t      3 others
ORIGIN

  Query Match            84.8%;  Score 17.8;  DB 10;  Length 466;
  Best Local Similarity  90.5%;  Pred. No. 4.5e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||| |||||||| |||
Db    329 GCTCTGACCACCTCAGGCAGG 349


RESULT 12
AV913575
LOCUS       AV913575                467 bp    mRNA    linear   EST 18-JAN-2002
DEFINITION  AV913575 K. Sato unpublished cDNA library, cv. Haruna Nijo
            germination shoots Hordeum vulgare subsp. vulgare cDNA clone
            bags22j03 5', mRNA sequence.
ACCESSION   AV913575
VERSION     AV913575.1  GI:18209352
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare.
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 467)
  AUTHORS   Sato,K., Saisho,D. and Takeda,K.
  TITLE     Barley EST sequencing project in NIG and Okayama Univ
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .467
                     /organism="Hordeum vulgare subsp. vulgare"
                     /cultivar="Haruna Nijo"
                     /db_xref="taxon:112509"
                     /clone="bags22j03"
                     /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
                     Nijo germination shoots"
                     /tissue_type="shoots"
                     /dev_stage="germination"
BASE COUNT       87 a    153 c    133 g     94 t
ORIGIN

  Query Match            84.8%;  Score 17.8;  DB 9;  Length 467;
  Best Local Similarity  90.5%;  Pred. No. 4.5e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||| |||||||| |||
Db    232 GCTCTGACCACCTCAGGCAGG 252


RESULT 13
BG263010
LOCUS       BG263010                470 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  WHE0939_G07_N13ZS Wheat 5-15 DAP spike cDNA library Triticum
            aestivum cDNA clone WHE0939_G07_N13, mRNA sequence.
ACCESSION   BG263010
VERSION     BG263010.1  GI:12864715
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 470)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - 5-15 DAP spike cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1. .470
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE0939_G07_N13"
                     /clone_lib="Wheat 5-15 DAP spike cDNA library"
                     /tissue_type="Spike"
                     /dev_stage="Adult plant"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
                     greenhouse. Spikes at 5, 10 and 15 DAP were  harvested,
                     total RNA and poly(A) RNA were prepared, a cDNA library
                     was  made, and the cDNA clones were in vivo excised to
                     give pBluescript  phagemids in the TJ Close lab (Choi,
                     Close, Fenton) at the University  of California,
                     Riverside. Plasmid DNA preparations and DNA sequencing
                     were performed in the OD Anderson lab (all other authors
                     )."
BASE COUNT       79 a    173 c    138 g     80 t
ORIGIN

  Query Match            84.8%;  Score 17.8;  DB 10;  Length 470;
  Best Local Similarity  90.5%;  Pred. No. 4.5e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTCTGACAACCTCAGGAAGG 21
          |||||||| |||||||| |||
Db    376 GCTCTGACCACCTCAGGCAGG 396


RESULT 14
BF630175
LOCUS       BF630175                492 bp    mRNA    linear   EST 22-OCT-2001
DEFINITION  HVSMEb0008D07f Hordeum vulgare seedling shoot EST library
            HVcDNA0002 (Dehydration stress) Hordeum vulgare cDNA clone
            HVSMEb0008D07f, mRNA sequence.
ACCESSION   BF630175
VERSION     BF630175.2  GI:13090694
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 492)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
            ,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton
```

```
          ,R.D., Oates,R. and Main,D.
  TITLE   Development of a genetically and physically anchored EST resource
          for barley genomics: Morex drought-stressed seedling shoot cDNA
          library
  JOURNAL Unpublished (2001)
COMMENT   On Dec 19, 2000 this sequence version replaced gi:11894333.
          Contact: Wing RA
          Clemson University Genomics Institute
          Clemson University
          100 Jordan Hall, Clemson, SC 29634, USA
          Tel: 864 656 7288
          Fax: 864 656 4293
          Email: rwing@clemson.edu
          Total hq bases = 418
          Seq primer: AATTAACCCTCACTAAAGGG
          High quality sequence stop: 482.
FEATURES             Location/Qualifiers
     source          1. .492
                     /organism="Hordeum vulgare"
                     /cultivar="Morex"
                     /db_xref="taxon:4513"
                     /clone="HVSMEb0008D07f"
                     /clone_lib="Hordeum vulgare seedling shoot EST library
                     HVcDNA0002 (Dehydration stress)"
                     /tissue_type="Seedling shoot"
                     /lab_host="TJC121"
                     /note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1;
                     Seeds were surface sterilized then germinated under axenic
                     conditions in the dark at room temperature on filter paper
                     with water, nystatin and cefotaxime in covered
                     crystallization dishes. Five-day old seedlings were
                     incubated at 90% RH for 24 hr. Shoots were then harvested,
                     total RNA was prepared, poly(A) RNA was purified, one
                     primary unamplified cDNA library was made, 600000 pfu were
                     in vivo excised to give pBluescript SK(-) cDNA phagemids.
                     These steps were performed in the TJ Close laboratory at
                     the University of California, Riverside (Choi, Close,
                     Fenton). Phagemids were plated and picked at the Clemson
                     University Genomics Institute (CUGI) (Begum, Palmer,
                     Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
                     sequencing and sequence analysis were performed at CUGI
                     (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
                     The sequence has been trimmed to remove vector sequence
                     and contains a minimum of 100 bases of phred value 20 or
                     above. For more details on library preparation and
                     sequence analysis see
                     http://www.genome.clemson.edu/projects/barley. To order
                     this clone see http://www.genome.clemson.edu/orders Also
                     see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                     Genetically and physically anchored EST resources for
                     barley genomics. Barley Genetics Newsletter 31:29-30.
                     (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT       82 a    201 c    137 g     72 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 10;  Length 492;
  Best Local Similarity   90.5%;  Pred. No. 4.5e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           ||||||||  |||||||| |||
Db     471 GCTCTGACCACCTCAGGCAGG 491


RESULT 15
BE406885
LOCUS       BE406885                 499 bp    mRNA    linear   EST 21-JUL-2000
DEFINITION  WHE0433_d05_g09zS Wheat etiolated seedling root cDNA library
            Triticum aestivum cDNA clone WHE0433_d05_g09, mRNA sequence.
ACCESSION   BE406885
VERSION     BE406885.1  GI:9366353
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 499)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Strategene SK primer.
FEATURES             Location/Qualifiers
     source          1. .499
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE0433_d05_g09"
                     /clone_lib="Wheat etiolated seedling root cDNA library"
                     /tissue_type="Root"
                     /dev_stage="Five day old etiolated seedling"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized
                     , germinated and grown aseptically in the  dark at room
                     temperature on filter paper with water, nystatin and
                     cefotaxime in covered crystallization dishes. Roots were
                     harvested.  The tissue, total RNA, and poly(A) RNA were
                     prepared, a cDNA library  was made, and the cDNA clones
                     were in vivo excised to give pBluescript  phagemids in the
                     TJ Close lab (Choi, Close, Fenton) at the University  of
                     California, Riverside.  Plasmid DNA preparations and DNA
                     sequencing  were performed in the OD Anderson lab (all
                     other authors)."
BASE COUNT       96 a    161 c    132 g    110 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 10;  Length 499;
  Best Local Similarity   90.5%;  Pred. No. 4.6e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GCTCTGACAACCTCAGGAAGG 21
           ||||||||  |||||||| |||
Db     181 GCTCTGACCACCTCAGGCAGG 201


Search completed: November  2, 2002, 06:42:11
Job time : 36.772 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:08:53 ; Search time 5094.62 Seconds
(without alignments)
10293.594 Million cell updates/sec

Title: US-09-981-606-1
Perfect score: 2506
Sequence: 1 atgggcccgcgagccaggcc.....ttgtattgtaaaaaaaaaaa 2506

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2506	100.0	2727	9	HSU60319	U60319 Homo sapien
2	1468	58.6	1885	9	AF144242	AF144242 Homo sapi
3	1219	48.6	1440	6	AR117793	AR117793 Sequence
4	1219	48.6	1440	6	AR149463	AR149463 Sequence
5	1217.4	48.6	1440	6	AR117794	AR117794 Sequence
6	1217.4	48.6	1440	6	AR117795	AR117795 Sequence
7	1217.4	48.6	1440	6	AR149464	AR149464 Sequence
8	1217.4	48.6	1440	6	AR149465	AR149465 Sequence
9	1215.8	48.5	1440	6	AR117796	AR117796 Sequence
10	1215.8	48.5	1440	6	AR149466	AR149466 Sequence
11	1178	47.0	1200	9	AF115265	AF115265 Homo sapi
12	1101.6	44.0	2332	4	AY007542	AY007542 Diceros b
13	1058	42.2	1479	9	HSHLAA6	Y09799 H.sapiens H
14	1051.6	42.0	10825	6	AR117789	AR117789 Sequence
15	1051.6	42.0	10825	6	AR117790	AR117790 Sequence
16	1051.6	42.0	10825	6	AR117791	AR117791 Sequence
17	1051.6	42.0	10825	6	AR117792	AR117792 Sequence
18	1051.6	42.0	10825	6	AR149459	AR149459 Sequence
19	1051.6	42.0	10825	6	AR149460	AR149460 Sequence
20	1051.6	42.0	10825	6	AR149461	AR149461 Sequence
21	1051.6	42.0	10825	6	AR149462	AR149462 Sequence
22	1051.6	42.0	12146	9	HSHFE	Z92910 Homo sapien
23	1051.6	42.0	101099	9	AL353759	AL353759 Human DNA
c 24	1051.6	42.0	193752	2	AL359892	AL359892 Homo sapi
25	1051.6	42.0	246240	6	AR036572	AR036572 Sequence
26	1051.6	42.0	246240	6	AR036573	AR036573 Sequence
27	1051.6	42.0	246240	6	AR036574	AR036574 Sequence
c 28	1051.6	42.0	246282	9	HSU91328	U91328 Human hered
29	1044	41.7	1280	9	HSA249335	AJ249335 Homo sapi
30	957	38.2	1045	9	AF079407	AF079407 Homo sapi
31	876.8	35.0	1320	4	AY007541	AY007541 Ceratothe
32	875.2	34.9	1320	4	AY007543	AY007543 Dicerorhi
33	875	34.9	1319	4	AY007544	AY007544 Rhinocero
34	849.4	33.9	1085	9	HSA249336	AJ249336 Homo sapi
35	711.4	28.4	823	9	AF079408	AF079408 Homo sapi
36	676.2	27.0	794	9	AF144238	AF144238 Homo sapi
37	654.8	26.1	1706	10	RNHERHAEM	AJ001517 Rattus no
38	636.4	25.4	1529	10	MMU66849	U66849 Mus musculu
39	627	25.0	1073	9	HSA249337	AJ249337 Homo sapi
40	617.4	24.6	781	9	AF079409	AF079409 Homo sapi
41	596	23.8	2115	9	HSA404378	AJ404378 Homo sapi
42	573	22.9	809	9	HSA250635	AJ250635 Homo sapi
43	508.4	20.3	987	9	AF150664	AF150664 Homo sapi
44	486.8	19.4	686	9	AF144240	AF144240 Homo sapi
45	478.2	19.1	789	10	AF008587	AF008587 Rattus no

ALIGNMENTS

RESULT 1
HSU60319
LOCUS HSU60319 2727 bp mRNA linear PRI 29-OCT-1997
DEFINITION Homo sapiens haemochromatosis protein (HLA-H) mRNA, complete cds.
ACCESSION U60319
VERSION U60319.1 GI:1469789
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2727)
AUTHORS Feder,J.N., Gnirke,A., Thomas,W., Tsuchihashi,Z., Ruddy,D.A., Basava,A., Dormishian,F., Domingo,R., Ellis,M.C., Fullan,A., Hinton,L.M., Jones,N.L., Kimmel,B.E., Kronmal,G.S., Lauer,P., Lee,V.K., Loeb,D.B., Mapa,F., McClelland,E., Meyer,N.C., Mintier,G.A., Moeller,N., Moore,T., Morkang,E., Prass,C.E., Quintana,L., Stranes,S.M., Schatzman,R.C., Brunke,K.J., Drayna,D.T., Risch,N.J., Bacon,B.R. and Wolff,R.K.
TITLE A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis
JOURNAL Nature Genet. 13 (4), 399-408 (1996)
MEDLINE 96331279

```
REFERENCE   2  (bases 1 to 2727)
  AUTHORS   Feder,J.N., Gnirke,A., Thomas,W., Tsuchihashi,Z., Ruddy,D.A.,
            Basava,A., Dormishian,F., Domingo,R., Ellis,M.C., Fullan,A.,
            Hinton,L.M., Jones,N.L., Kimmel,B.E., Kronmal,G.S., Lauer,P.,
            Lee,V.K., Loeb,D.B., Mapa,F., McClelland,E., Meyer,N.C.,
            Mintier,G.A., Moeller,N., Moore,T., Morkang,E., Prass,C.E.,
            Quintana,L., Stranes,S.M., Schatzman,R.C., Brunke,K.J.,
            Drayna,D.T., Risch,N.J., Bacon,B.R. and Wolff,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-1996) Mercator Genetics, 4040 Campbell Ave.,
            Menlo Park, CA 94025, USA
FEATURES             Location/Qualifiers
     source          1. .2727
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.3"
     gene            1. .2727
                     /gene="HLA-H"
                     /note="HFE"
     CDS             222. .1268
                     /gene="HLA-H"
                     /codon_start=1
                     /product="haemochromatosis protein"
                     /protein_id="AAC51823.1"
                     /db_xref="GI:1469790"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                     SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
                     WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                     RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTS
                     SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGE
                     EQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQG
                     SRGAMGHYVLAERE"
BASE COUNT      702 a    606 c    660 g    759 t
ORIGIN

  Query Match           100.0%;  Score 2506;  DB 9;  Length 2727;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2506;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy    1201 CCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1422 CCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTG 1481

Qy    1261 TGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1482 TGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCA 1541

Qy    1321 TTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1542 TTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAG 1601

Qy    1381 GACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1602 GACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAAC 1661

Qy    1441 CCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACATGTATCTATGCATTTTCT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1662 CCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACATGTATCTATGCATTTTCT 1721

Qy    1501 GGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1722 GGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCA 1781

Qy    1561 CCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCACAGCTATGAAGGCTGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1782 CCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCACAGCTATGAAGGCTGT 1841

Qy   1621 GCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTGGGTA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1842 GCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTGGGTA 1901

Qy   1681 GTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGT 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1902 GTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGT 1961

Qy   1741 GTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1962 GTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCAT 2021

Qy   1801 TTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2022 TTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTC 2081

Qy   1861 TCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGCACTTACTTCGTGTCCGA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2082 TCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGCACTTACTTCGTGTCCGA 2141

Qy   1921 CTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTTACAATAATTCTATGAG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2142 CTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTTACAATAATTCTATGAG 2201

Qy   1981 ATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGCACGG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2202 ATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGCACGG 2261

Qy   2041 TGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2262 TGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCC 2321

Qy   2101 GTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2322 GTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATG 2381

Qy   2161 CACTTGGCTGCATAAATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATA 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2382 CACTTGGCTGCATAAATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATA 2441

Qy   2221 CCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCT 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2442 CCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCT 2501

Qy   2281 CTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2502 CTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATG 2561

Qy   2341 TGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2562 TGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATT 2621

Qy   2401 AGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2622 AGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAA 2681

Qy   2461 AATGCATATACTTTAATAAATGTACATTGTATTGTAAAAAAAAAAA 2506
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   2682 AATGCATATACTTTAATAAATGTACATTGTATTGTAAAAAAAAAAA 2727
```

RESULT 2
AF144242

```
LOCUS       AF144242               1885 bp    mRNA    linear   PRI 07-MAY-2001
DEFINITION  Homo sapiens hemochromatosis splice variant delE3 mRNA, complete
            cds.
ACCESSION   AF144242
VERSION     AF144242.1  GI:11094324
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1885)
  AUTHORS   Thenie,A., Orhant,M., Gicquel,I., Fergelot,P., Le Gall,J.Y.,
            David,V. and Mosser,J.
  TITLE     The HFE gene undergoes alternate splicing processes
  JOURNAL   Blood Cells Mol. Dis. 26 (2), 155-162 (2000)
  MEDLINE   20448010
   PUBMED   11001625
REFERENCE   2  (bases 1 to 1885)
  AUTHORS   Thenie,A., Orhant,M., Gicquel,I. and Mosser,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-APR-1999) Faculte de Medecine, UPR41 CNRS, 2 Avenue
            du Pr. Leon Bernard, Rennes Cedex 35043, France
FEATURES             Location/Qualifiers
     source          1. .1885
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22"
     gene            1. .1885
                     /gene="HFE"
     CDS             124. .894
                     /gene="HFE"
                     /codon_start=1
                     /product="hemochromatosis splice variant delE3"
                     /protein_id="AAG29577.1"
                     /db_xref="GI:11094325"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                     SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
                     WTIMENHNHSKVPPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFE
                     PKDALPNGDGTYQGWITLAVPPGEEQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVI
                     SGIAVFVVILFIGILFIILRKRQGSRGAVGHYVLAERE"
BASE COUNT      453 a    442 c    458 g    532 t
ORIGIN
```

```
  Query Match           58.6%;  Score 1468;  DB 9;  Length 1885;
  Best Local Similarity  86.2%;  Pred. No. 0;
  Matches 1757;  Conservative    0;  Mismatches    5;  Indels  276;  Gaps    1;

Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 183

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 243

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 303

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 363

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 423

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||
Db    424 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAG---------------------- 462

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420

Db    463 ---------------------------------------------------------- 462

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480

Db    463 ---------------------------------------------------------- 462

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
```

```
Db    463 ------------------------------------------------------------ 462

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600

Db    463 ------------------------------------------------------------ 462

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    463 --------------GTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 507

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    508 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 567

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    568 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGCATTGCCCAATGGG 627

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    628 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 687

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    688 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 747

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    748 CCGTCTGGCACCCTAGTCATTGGAGTCATTAGTGGAATTGCTGTTTTTGTCGTCATCTTG 807

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    808 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCGTGGGG 867

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    868 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 927

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    928 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 987

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    988 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1047

Qy   1201 CCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1048 CCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTG 1107

Qy   1261 TGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1108 TGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCA 1167

Qy   1321 TTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1168 TTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAG 1227

Qy   1381 GACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1228 GACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAAC 1287

Qy   1441 CCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACATGTATCTATGCATTTTCT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1288 CCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACATGTATCTATGCATTTTCT 1347

Qy   1501 GGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1348 GGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCA 1407

Qy   1561 CCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCACAGCTATGAAGGCTGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1408 CCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCACAGCTATGAAGGCTGT 1467

Qy   1621 GCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTGGGTA 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1468 ACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTGGGTA 1527

Qy   1681 GTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGT 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1528 GTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGT 1587

Qy   1741 GTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCAT 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   1588 GTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCCTCAT 1647

Qy   1801 TTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1648 TTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTC 1707

Qy   1861 TCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGCACTTACTTCGTGTCCGA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1708 TCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGCACTTACTTCGTGTCCGA 1767

Qy   1921 CTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTTACAATAATTCTATGAG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1768 CTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTTACAATAATTCTATGAG 1827

Qy   1981 ATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGCAC 2038
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1828 ATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGCAC 1885
```

```
RESULT 3
AR117793
LOCUS       AR117793               1440 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 9 from patent US 6140305.
ACCESSION   AR117793
VERSION     AR117793.1  GI:14098699
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 9 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      347 a    355 c    407 g    331 t
ORIGIN

  Query Match             48.6%;  Score 1219;  DB 6;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-303;
  Matches 1219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
```

```
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
Db   1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
Db   1422 CCTCAAAAAGATTTCCCCA 1440

RESULT 4
AR149463
LOCUS       AR149463               1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 9 from patent US 6228594.
ACCESSION   AR149463
VERSION     AR149463.1  GI:15114054
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 9 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      347 a    355 c    407 g    331 t
ORIGIN

  Query Match             48.6%;  Score 1219;  DB 6;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-303;
  Matches 1219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
Db    402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001
```

```
Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
Db   1422 CCTCAAAAAGATTTCCCCA 1440
```

RESULT 5
AR117794
LOCUS AR117794 1440 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 10 from patent US 6140305.
ACCESSION AR117794
VERSION AR117794.1 GI:14098700
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 10 31-OCT-2000;
FEATURES Location/Qualifiers
source 1. .1440
/organism="unknown"
BASE COUNT 348 a 355 c 406 g 331 t
ORIGIN

Query Match 48.6%; Score 1217.4; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 8.6e-303;
Matches 1218; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461
```

```
Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
Db   1422 CCTCAAAAAGATTTCCCCA 1440
```

RESULT 6
AR117795
LOCUS AR117795 1440 bp DNA linear PAT 16-MAY-2001

```
DEFINITION  Sequence 11 from patent US 6140305.
ACCESSION   AR117795
VERSION     AR117795.1  GI:14098701
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 11 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      347 a    354 c    408 g    331 t
ORIGIN

  Query Match           48.6%;  Score 1217.4;  DB 6;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 8.6e-303;
  Matches 1218;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
Db   1422 CCTCAAAAAGATTTCCCCA 1440


RESULT 7
AR149464
LOCUS       AR149464               1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 10 from patent US 6228594.
ACCESSION   AR149464
VERSION     AR149464.1  GI:15114055
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 10 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      348 a    355 c    406 g    331 t
ORIGIN

  Query Match           48.6%;  Score 1217.4;  DB 6;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 8.6e-303;
  Matches 1218;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
```

```
Db    402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
Db   1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
Db   1422 CCTCAAAAAGATTTCCCCA 1440

RESULT 8
```

```
AR149465
LOCUS       AR149465               1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 11 from patent US 6228594.
ACCESSION   AR149465
VERSION     AR149465.1  GI:15114056
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 11 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      347 a    354 c    408 g    331 t
ORIGIN

  Query Match             48.6%;  Score 1217.4;  DB 6;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 8.6e-303;
  Matches 1218;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
Db    402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941
```

```
Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
Db   1422 CCTCAAAAAGATTTCCCCA 1440
```

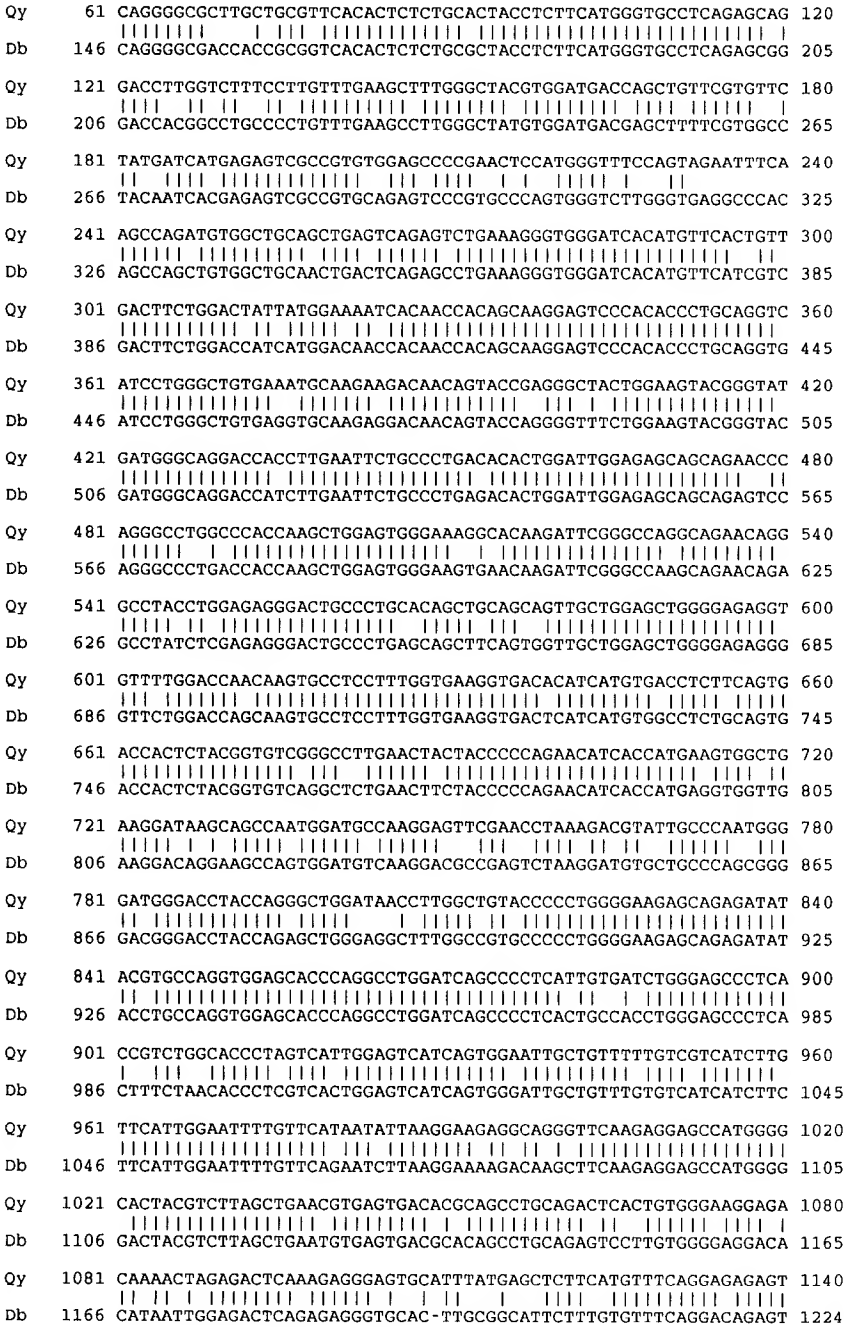
RESULT 9
AR117796
LOCUS AR117796 1440 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 12 from patent US 6140305.
ACCESSION AR117796
VERSION AR117796.1 GI:14098702
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 12 31-OCT-2000;
FEATURES Location/Qualifiers
source 1. .1440
/organism="unknown"
BASE COUNT 348 a 354 c 407 g 331 t
ORIGIN

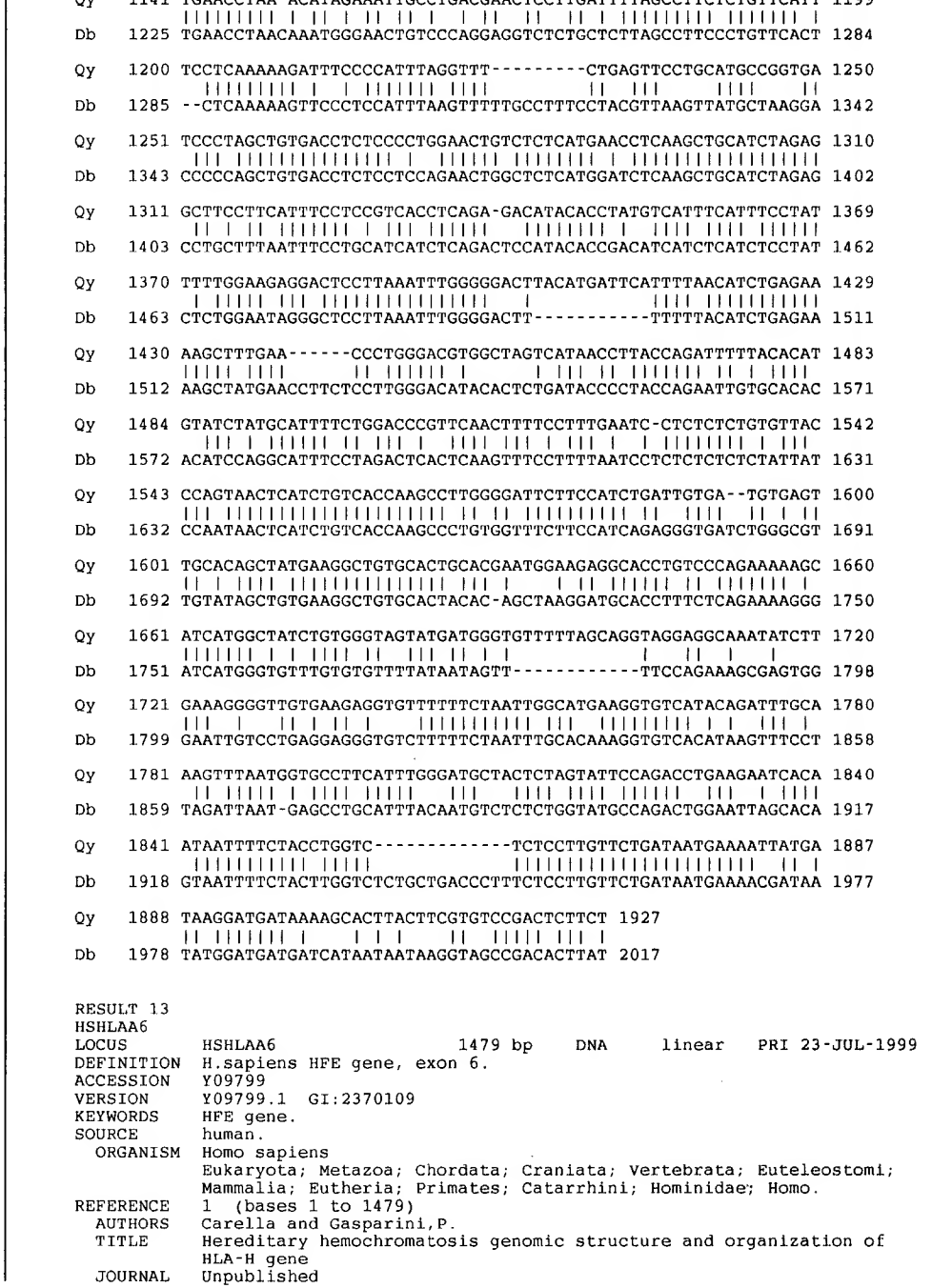
Query Match 48.5%; Score 1215.8; DB 6; Length 1440;
Best Local Similarity 99.8%; Pred. No. 2.2e-302;
Matches 1217; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
Db   1422 CCTCAAAAAGATTTCCCCA 1440
```

```
RESULT 10
AR149466
LOCUS       AR149466               1440 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 12 from patent US 6228594.
ACCESSION   AR149466
VERSION     AR149466.1  GI:15114057
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Method for determining the presence or absence of a hereditary
            hemochromatosis gene mutation
  JOURNAL   Patent: US 6228594-A 12 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="unknown"
BASE COUNT      348 a    354 c    407 g    331 t
ORIGIN

  Query Match             48.5%;  Score 1215.8;  DB 6;  Length 1440;
  Best Local Similarity   99.8%;  Pred. No. 2.2e-302;
  Matches 1217;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy    1201 CCTCAAAAAGATTTCCCCA 1219
           ||||||||||||||||||
Db    1422 CCTCAAAAAGATTTCCCCA 1440


RESULT 11
AF115265
LOCUS       AF115265               1200 bp    mRNA    linear   PRI 07-MAY-2001
DEFINITION  Homo sapiens hemochromatosis termination variant terE6 (HFE) mRNA,
            complete cds.
ACCESSION   AF115265
VERSION     AF115265.1  GI:11094314
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1200)
  AUTHORS   Thenie,A., Orhant,M., Gicquel,I., Fergelot,P., Le Gall,J.Y.,
            David,V. and Mosser,J.
  TITLE     The HFE gene undergoes alternate splicing processes
  JOURNAL   Blood Cells Mol. Dis. 26 (2), 155-162 (2000)
  MEDLINE   20448010
   PUBMED   11001625
REFERENCE   2  (bases 1 to 1200)
  AUTHORS   Thenie,A., Orhant,M. and Mosser,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-1998) UPR 41 CNRS, Faculte de Medecine, 2, av du
            Pr. Bernard, Rennes 35043, France
FEATURES             Location/Qualifiers
     source          1. .1200
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p22"
     gene            1. .1200
                     /gene="HFE"
     CDS             12. .1058
                     /gene="HFE"
```

```
                     /codon_start=1
                     /product="hemochromatosis termination variant terE6"
                     /protein_id="AAG29572.1"
                     /db_xref="GI:11094315"
                     /translation="MGPRARPALLLLMLLQTAVLQGRLLRSHSLHYLFMGASEQDLGL
                     SLFEALGYVDDQLFVFYDHESRRVEPRTPWVSSRISSQMWLQLSQSLKGWDHMFTVDF
                     WTIMENHNHSKESHTLQVILGCEMQEDNSTEGYWKYGYDGQDHLEFCPDTLDWRAAEP
                     RAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLLELGRGVLDQQVPPLVKVTHHVTS
                     SVTTLRCRALNYYPQNITMKWLKDKQPMDAKEFEPKDVLPNGDGTYQGWITLAVPPGE
                     EQRYTCQVEHPGLDQPLIVIWEPSPSGTLVIGVISGIAVFVVILFIGILFIILRKRQG
                     SRGAMGHYVLAERE"
BASE COUNT      298 a    290 c    346 g    266 t
ORIGIN

  Query Match           47.0%;  Score 1178;  DB 9;  Length 1200;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-292;
  Matches 1178;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 71

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 131

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     132 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 191

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     192 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 251

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     252 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 311

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     312 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 371

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     372 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 431

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     432 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 491

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     492 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 551

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     552 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 611

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     612 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 671

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     672 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 731

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     732 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 791

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     792 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 851

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     852 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 911

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     912 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 971

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     972 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1031

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1032 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1091

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1092 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1151

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGA 1178
             ||||||||||||||||||||||||||||||||||||||
Db    1152 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGA 1189


RESULT 12
AY007542
LOCUS       AY007542               2332 bp    mRNA    linear   MAM 22-OCT-2000
DEFINITION  Diceros bicornis HFE protein mRNA, complete cds.
ACCESSION   AY007542
VERSION     AY007542.1  GI:10945689
KEYWORDS    .
SOURCE      black rhinoceros.
  ORGANISM  Diceros bicornis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
REFERENCE   1  (bases 1 to 2332)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Rhinoceros HFE Polymorphisms
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2332)
  AUTHORS   West,C.J., Worley,M. and Beutler,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Molecular and Experimental Medicine, The
            Scripps Research Institute, 10550 North Torrey Pines Road, La
            Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .2332
                     /organism="Diceros bicornis"
                     /db_xref="taxon:9805"
     CDS             86. .1132
                     /codon_start=1
                     /product="HFE protein"
                     /protein_id="AAG23702.1"
                     /db_xref="GI:10945690"
                     /translation="MGPRARPALFFLILLRTVAAQGRPPRSHSLRYLFMGASERDHGL
                     PLFEALGYVDDELFVAYNHESRRAESRAQWVLGEAHSQLWLQLTQSLKGWDHMFIVDF
                     WTIMDNHNHSKESHTLQVILGCEVQEDNSTRGFWKYGYDGQDHLEFCPETLDWRAAES
                     RALTTKLEWEVNKIRAKQNRAYLERDCPEQLQWLLELGRGVLDQQVPPLVKVTHHVAS
                     AVTTLRCQALNFYPQNITMRWLKDRKPVDVKDAESKDVLPSGDGTYQSWEALAVPPGE
                     EQRYTCQVEHPGLDQPLTATWEPSLSNTLVTGVISGIAVCVIIFFIGILFRILRKRQA
                     SRGAMGDYVLAECE"
BASE COUNT      581 a    592 c    561 g    598 t
ORIGIN

  Query Match           44.0%;  Score 1101.6;  DB 4;  Length 2332;
  Best Local Similarity  77.6%;  Pred. No. 6.9e-273;
  Matches 1520;  Conservative    0;  Mismatches  379;  Indels   61;  Gaps   13;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
             |||||||||||||| | || |||||  || |||||| ||  ||| ||||| || |  |
Db      86 ATGGGCCCGCGAGCCCGACCCGCGCTCTTCTTCCTGATCCTCCTGCGGACCGTGGCCGCG 145
```





RESULT 13
HSHLAA6
LOCUS HSHLAA6 1479 bp DNA linear PRI 23-JUL-1999
DEFINITION H.sapiens HFE gene, exon 6.
ACCESSION Y09799
VERSION Y09799.1 GI:2370109
KEYWORDS HFE gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1479)
AUTHORS Carella and Gasparini,P.
TITLE Hereditary hemochromatosis genomic structure and organization of
HLA-H gene
JOURNAL Unpublished

```
REFERENCE   2  (bases 1 to 1479)
  AUTHORS   Gasparini,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1996) P. Gasparini, Servizio de Genetica Medica -
            IRCCS, 'Ospedale CSS', Via Cappuccini, 71013 S Giovanni, Rotondo
            (FG), ITALY
COMMENT     Related sequence: U60319.
FEATURES             Location/Qualifiers
     source          1. .1479
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="31H6"
                     /map="6p22"
     misc_feature    349. .389
                     /gene="HFE"
                     /label=stop
                     /usedin=Y09801:hfe_cds
     gene            349. .1425
                     /gene="HFE"
     exon            349. .1425
                     /gene="HFE"
                     /label=ex6
                     /usedin=Y09801:hfe_mrna
BASE COUNT      397 a    298 c    358 g    426 t
ORIGIN

  Query Match             42.2%;  Score 1058;  DB 9;  Length 1479;
  Best Local Similarity   99.1%;  Pred. No. 1.1e-261;
  Matches 1064;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    346 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 405

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     406 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 465

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     466 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 525

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     526 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 585

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     586 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 645

Qy   1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     646 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 705

Qy   1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     706 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 765

Qy   1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     766 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 825

Qy   1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     826 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 885

Qy   1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     886 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 945

Qy   1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     946 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1005

Qy   1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1006 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1065

Qy   1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1066 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1125

Qy   1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1126 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1185

Qy   1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1186 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1245

Qy   1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1246 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1305

Qy   1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1306 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 1365

Qy   2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
          ||||||||||||||||||||||||||||||||||||||||     || ||   ||
Db    1366 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGCACTTTGGGAGGCCA 1419


RESULT 14
AR117789
LOCUS       AR117789               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6140305.
ACCESSION   AR117789
VERSION     AR117789.1  GI:14098695
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 1 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
BASE COUNT     2998 a   2253 c   2648 g   2926 t
ORIGIN

  Query Match             42.0%;  Score 1051.6;  DB 6;  Length 10825;
  Best Local Similarity   98.7%;  Pred. No. 7.4e-260;
  Matches 1060;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7104 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 7163

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7164 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 7223

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7224 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 7283

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7284 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 7343

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    7344 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 7403

Qy    1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7404 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 7463

Qy    1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7464 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 7523

Qy    1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7524 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 7583

Qy    1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7584 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 7643

Qy    1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7644 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 7703

Qy    1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    7704 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 7763

Qy    1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7764 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 7823

Qy    1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7824 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 7883

Qy    1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7884 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 7943

Qy    1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7944 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 8003

Qy    1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8004 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 8063

Qy    1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8064 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 8123

Qy    2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
           ||||||||||||||||||||| ||||||| ||||||     || ||   ||
Db    8124 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 8177
```

```
RESULT 15
AR117790
LOCUS       AR117790               10825 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 3 from patent US 6140305.
ACCESSION   AR117790
VERSION     AR117790.1  GI:14098696
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 10825)
  AUTHORS   Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
            Tsuchihashi,Z. and Wolff,R.K.
  TITLE     Hereditary hemochromatosis gene products
  JOURNAL   Patent: US 6140305-A 3 31-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .10825
                     /organism="unknown"
```

```
BASE COUNT     2999 a   2253 c   2647 g   2926 t
ORIGIN

  Query Match           42.0%;  Score 1051.6;  DB 6;  Length 10825;
  Best Local Similarity  98.7%;  Pred. No. 7.4e-260;
  Matches 1060;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7104 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 7163

Qy    1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7164 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 7223

Qy    1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7224 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 7283

Qy    1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7284 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 7343

Qy    1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7344 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 7403

Qy    1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7404 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 7463

Qy    1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7464 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 7523

Qy    1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7524 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 7583

Qy    1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7584 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 7643

Qy    1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7644 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 7703

Qy    1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    7704 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 7763

Qy    1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7764 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 7823

Qy    1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7824 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 7883

Qy    1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7884 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 7943

Qy    1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7944 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 8003

Qy    1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8004 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 8063

Qy    1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    8064 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 8123

Qy    2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
           ||||||||||||||||||||||| |||||||  ||||||    || ||   ||
Db    8124 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 8177
```

Search completed: November 2, 2002, 05:38:07
Job time : 5138.62 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:05:23 ; Search time 510.445 Seconds
(without alignments)
8429.091 Million cell updates/sec

Title: US-09-981-606-1
Perfect score: 2506
Sequence: 1 atgggcccgcgagccaggcc.....ttgtattgtaaaaaaaaaaa 2506

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2506	100.0	2506	21	AAA96769	cDNA sequence enco
	2	2506	100.0	2727	19	AAV23525	Haemochromatosis g
	3	1219	48.6	1440	18	AAT96691	Hereditary haemoch
	4	1219	48.6	1440	22	AAC68429	Human hereditary h
	5	1217.4	48.6	1440	22	AAC68430	Human hereditary h
	6	1217.4	48.6	1440	22	AAC68431	Human hereditary h
	7	1215.8	48.5	1440	22	AAC68432	Human hereditary h
	8	1051.6	42.0	5749	22	AAL36747	Human musculoskele
	9	1051.6	42.0	10825	18	AAT96690	Hereditary haemoch
	10	1051.6	42.0	10825	22	AAC68425	Human hereditary h
	11	1051.6	42.0	10825	22	AAC68426	Human hereditary h
	12	1051.6	42.0	10825	22	AAC68427	Human hereditary h
	13	1051.6	42.0	10825	22	AAC68428	Human hereditary h
	14	1051.6	42.0	12146	21	AAA96794	Genomic DNA of a h
c	15	1048.4	41.8	237326	19	AAV57903	Hereditary haemoch
	16	503.2	20.1	596	22	AAI63897	Human polynucleoti
c	17	456.2	18.2	235033	19	AAV57926	Hereditary haemoch
	18	432	17.2	1712	22	AAL36748	Human musculoskele
	19	304.4	12.1	306	22	AAL36750	Human musculoskele
	20	280	11.2	359	20	AAX16055	Hereditary hemochr
	21	280	11.2	517	22	AAC68440	Human hereditary h
	22	278.4	11.1	517	22	AAC68441	Human hereditary h
	23	175.2	7.0	1112	21	AAA48668	cDNA encoding chic
	24	173.6	6.9	1230	21	AAA48673	cDNA encoding chic
	25	170.4	6.8	1195	21	AAA48671	cDNA encoding chic
	26	168.4	6.7	1197	21	AAA48672	cDNA encoding chic
	27	168.4	6.7	1262	21	AAA48691	Consensus DNA of c
	28	166.8	6.7	1173	21	AAC78071	Human cancer assoc
	29	166.8	6.7	1173	22	AAF72767	Human prostate can
	30	163.6	6.5	1145	21	AAA48667	cDNA encoding chic
	31	162	6.5	1230	21	AAA48669	cDNA encoding chic
	32	162	6.5	1284	9	AAN80603	Probe F10 of Major
	33	160.4	6.4	1230	21	AAA48665	cDNA encoding chic
	34	160.4	6.4	1230	21	AAA48670	cDNA encoding chic
	35	158.8	6.3	1230	21	AAA48666	cDNA encoding chic
	36	157.2	6.3	1230	21	AAA48664	cDNA encoding chic
	37	145.2	5.8	1554	22	AAI93004	Human polynucleoti
	38	144.8	5.8	1101	12	AAQ12116	HLA-C exon Cb-1.
	39	143.6	5.7	1567	22	AAH98676	Human EST-derived
	40	143.4	5.7	1073	22	AAH42223	Nucleotide sequenc
	41	142.8	5.7	1098	22	AAS07697	Human cDNA encodin
	42	141.6	5.7	1101	12	AAQ12117	HLA-C exon Cb-2.
	43	141	5.6	2034	23	AAS90913	DNA encoding novel
	44	141	5.6	2037	23	AAS90740	DNA encoding novel
	45	140	5.6	4965	16	AAQ75973	pHLA-B7/beta-2 mic

ALIGNMENTS

```
RESULT 1
AAA96769
ID   AAA96769 standard; cDNA; 2506 BP.
XX
AC   AAA96769;
XX
DT   19-FEB-2001  (first entry)
XX
DE   cDNA sequence encoding a histocompatibility iron loading (HFE) protein.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1044
FT                   /*tag= a
FT                   /product= "histocompatibility iron loading (HFE) protein"
FT   sig_peptide     1..66
FT                   /*tag= b
FT   mutation        187
FT                   /*tag= c
FT                   /note= "if this base is mutated to G, then the
FT                        protein contains the mutation H63D"
FT   mutation        193
FT                   /*tag= d
FT                   /note= "if this base is mutated to T, then the
FT                        protein contains the mutation S65C"
FT   mutation        277
FT                   /*tag= e
```

```
FT                   /note= "if this base is mutated to C, then the
FT                          protein contains the mutation G93R"
FT   mutation        314
FT                   /*tag= f
FT                   /note= "if this base is mutated to C, then the
FT                          protein contains the mutation I105T, which
FT                          is associated with an iron overload disorder"
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
DR   P-PSDB; AAB19149.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Disclosure; Page 2-3; 55pp; English.
XX
CC   The present sequence encodes a human histocompatibility iron loading
CC   (HFE) protein. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence 2506 BP; 648 A; 552 C; 596 G; 710 T; 0 other;

  Query Match             100.0%;  Score 2506;  DB 21;  Length 2506;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2506;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200

Qy    1201 CCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTG 1260

Qy    1261 TGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 TGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCA 1320

Qy    1321 TTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 TTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAG 1380

Qy    1381 GACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAAC 1440

Qy    1441 CCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACATGTATCTATGCATTTTCT 1500
```

```
Db   1441 CCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACATGTATCTATGCATTTTCT 1500

Qy   1501 GGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCA 1560

Qy   1561 CCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCACAGCTATGAAGGCTGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCACAGCTATGAAGGCTGT 1620

Qy   1621 GCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTGGGTA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTGGGTA 1680

Qy   1681 GTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGT 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGT 1740

Qy   1741 GTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCAT 1800

Qy   1801 TTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTC 1860

Qy   1861 TCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGCACTTACTTCGTGTCCGA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGCACTTACTTCGTGTCCGA 1920

Qy   1921 CTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTTACAATAATTCTATGAG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTTACAATAATTCTATGAG 1980

Qy   1981 ATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGCACGG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGCACGG 2040

Qy   2041 TGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCC 2100

Qy   2101 GTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATG 2160

Qy   2161 CACTTGGCTGCATAAATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATA 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 CACTTGGCTGCATAAATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATA 2220

Qy   2221 CCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCT 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCT 2280

Qy   2281 CTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATG 2340

Qy   2341 TGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATT 2400

Qy   2401 AGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 AGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAA 2460

Qy   2461 AATGCATATACTTTAATAAATGTACATTGTATTGTAAAAAAAAAAA 2506
          ||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AATGCATATACTTTAATAAATGTACATTGTATTGTAAAAAAAAAAA 2506
```

RESULT 2

```
AAV23525
ID   AAV23525 standard; mRNA; 2727 BP.
XX
AC   AAV23525;
XX
DT   10-JUL-1998  (first entry)
XX
DE   Haemochromatosis gene.
XX
KW   Hereditary haemochromatosis; HC gene; HH identification; diagnosis;
KW   autosomal recessive disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO9807884-A1.
XX
PD   26-FEB-1998.
XX
PF   22-AUG-1997;   97WO-AU00539.
XX
PR   03-SEP-1996;   96AU-0002083.
PR   23-AUG-1996;   96AU-0001849.
XX
PA   (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX
PI   Busfield F,  Cullen LM,  Jazwinska EC,  Powell LW;
XX
DR   WPI; 1998-179064/16.
XX
PT   Detection of autosomal recessive disorder - particularly hereditary
PT   haemochromatosis, by detecting a mutation in the HC gene
XX
PS   Disclosure; Page -; 32pp; English.
XX
CC   This sequence represents the haemochromatosis (HC) gene. Mutations in
CC   this sequence are detected using the method of the invention. The method
CC   is for identifying an individual with hereditary haemochromatosis (HH) or
CC   a predisposition to develop HH or to genetically pass on HH to an
CC   offspring, comprising isolating a biological sample and amplifying a
CC   region of genomic DNA in the biological sample encompassing all or part
CC   of the DNA between markers D6S265 and D6S276, and detecting at least one
CC   homozygous or heterozygous mutation in a nucleotide within the region.
CC   The method can also be used for identifying an individual with an
CC   autosomal recessive disorder (ARD) or predisposition to develop and/or
CC   genetically pass on an ARD to an offspring, comprises isolating a
CC   biological sample from the individual and screening genomic DNA in the
CC   sample for the presence of a homozygous or heterozygous mutation in a
CC   gene, the normal function of which, is required to prevent progression of
CC   the disorder. The method(s) can be used to identify individuals that are
CC   homozygous or heterozygous (carriers) for the mutation causing the ARD.
CC   Especially the method is used to diagnose HH or predisposition to HH by
CC   detecting a Cys282Tyr substitution. Individuals homozygous for this
CC   mutation have HH and heterozygotes are potential carriers of the
CC   disease.
XX
SQ   Sequence 2727 BP; 702 A; 606 C; 660 G; 759 T; 0 other;

  Query Match             100.0%;  Score 2506;  DB 19;  Length 2727;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2506;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401
```


Qy 181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
Db 402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461
Qy 241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
Db 462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521
Qy 301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
Db 522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581
Qy 361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
Db 582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641
Qy 421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
Db 642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701
Qy 481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
Db 702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761
Qy 541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
Db 762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821
Qy 601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
Db 822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881
Qy 661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
Db 882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941
Qy 721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
Db 942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001
Qy 781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
Db 1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061
Qy 841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
Db 1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121
Qy 901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
Db 1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181
Qy 961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
Db 1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241
Qy 1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
Db 1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301
Qy 1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
Db 1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361
Qy 1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
Db 1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421
Qy 1201 CCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTG 1260
Db 1422 CCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTG 1481
Qy 1261 TGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCA 1320
Db 1482 TGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCA 1541
Qy 1321 TTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAG 1380
Db 1542 TTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAG 1601
Qy 1381 GACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAAC 1440
Db 1602 GACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAAC 1661
Qy 1441 CCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACATGTATCTATGCATTTTCT 1500
Db 1662 CCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACATGTATCTATGCATTTTCT 1721
Qy 1501 GGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCA 1560
Db 1722 GGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCA 1781
Qy 1561 CCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCACAGCTATGAAGGCTGT 1620
Db 1782 CCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCACAGCTATGAAGGCTGT 1841
Qy 1621 GCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTGGGTA 1680
Db 1842 GCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTGGGTA 1901
Qy 1681 GTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGT 1740
Db 1902 GTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGT 1961
Qy 1741 GTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCAT 1800
Db 1962 GTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCAT 2021
Qy 1801 TTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTC 1860
Db 2022 TTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTC 2081
Qy 1861 TCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGCACTTACTTCGTGTCCGA 1920
Db 2082 TCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGCACTTACTTCGTGTCCGA 2141
Qy 1921 CTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTTACAATAATTCTATGAG 1980
Db 2142 CTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTTACAATAATTCTATGAG 2201
Qy 1981 ATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGCACGG 2040
Db 2202 ATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGCACGG 2261
Qy 2041 TGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCC 2100
Db 2262 TGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCC 2321
Qy 2101 GTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATG 2160
Db 2322 GTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATG 2381
Qy 2161 CACTTGGCTGCATAAATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATA 2220
Db 2382 CACTTGGCTGCATAAATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATA 2441
Qy 2221 CCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCT 2280
Db 2442 CCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCT 2501
Qy 2281 CTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATG 2340
Db 2502 CTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATG 2561
Qy 2341 TGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATT 2400

```
Db    2562 TGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATT 2621

Qy    2401 AGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAA 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2622 AGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAA 2681

Qy    2461 AATGCATATACTTTAATAAATGTACATTGTATTGTAAAAAAAAAAA 2506
           |||||||||||||||||||||||||||||||||||||||||||||
Db    2682 AATGCATATACTTTAATAAATGTACATTGTATTGTAAAAAAAAAAA 2727


RESULT 3
AAT96691
ID   AAT96691 standard; cDNA; 1440 BP.
XX
AC   AAT96691;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene cDNA clone.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis;
KW   gene therapy; prenatal screening; human; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             222.1268
FT                   /*tag=  a
FT   mutation        408
FT                   /*tag=  g
FT                   /note= "C to G substitution (24d2 mutation)
FT                          results in His to Asp substitution"
FT   variation       414
FT                   /*tag=  h
FT                   /note= "A to T substitution (24d7 variant)
FT                          results in Ser to Cys substitution"
FT   mutation        1066
FT                   /*tag=  i
FT                   /note= "G to A substitution (24d1 mutation
FT                          associated with HH), results in Cys to
FT                          Tyr substitution"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US06254.
XX
PR   23-MAY-1996;   96US-0652265.
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
XX
PA   (MERC-) MERCATOR GENETICS INC.
XX
PI   Drayna DT,  Feder JN,  Gnirke A,  Ruddy D,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1997-512743/47.
DR   P-PSDB; AAW36499.
XX
PT   Hereditary haemochromatosis gene and variants - useful for diagnosis
PT   and treatment of hereditary haemochromatosis disease
XX
PS   Disclosure; Fig 4; 115pp; English.
XX
CC   This cDNA clone, designated cDNA24, is derived from human gene
CC   whose mutated form is associated with hereditary haemochromatosis
CC   (HH).  It was obtained from a directionally cloned plasmid-based
CC   cDNA library following identification of the HH locus in the HLA
CC   region of chromosome 6.  A single mutation (24d1) in the HH gene
CC   appears responsible for the majority of HH disease.  This comprises
CC   a G to A substitution that is present in 86% of affected
CC   chromosomes and in 4% of unaffected chromosomes.  It results in a
CC   Cys to Tyr substitution in the encoded protein (see AAW36499) at a
CC   critical disulphide bridge important for secondary structure.  The
CC   following are claimed: a 10825 bp genomic DNA sequence (I) (see
CC   AAT96690), the 1437 bp cDNA sequence (Ia) and their 24d1, 24d2 and
CC   24d7 variants; a cloning or expression vector; host cells; a
CC   peptide product chosen from the HH gene product, its variants
CC   (24d1, 24d2 and 24d7), or a peptide of at least 56 amino acid
CC   residues of these; an antibody produced using the peptide; a method
CC   to determine the presence or absence of the common HH gene
CC   mutation; an animal model for the HH disease; metal chelation
CC   agents, T-cell differentiation factors and therapeutic agents for
CC   the mitigation of injury due to oxidative process in vivo or
CC   mitigation of iron overload; a method for screening potential
CC   therapeutic agents for activity in connection with HH disease; an
CC   antisense oligonucleotide directed against a transcriptional
CC   product of a nucleic acid sequence as above; and oligonucleotides
CC   or pairs of oligonucleotides covering a range of nucleotides from
CC   (I), (Ia) or their variants, useful for detecting a polymorphism in
CC   the HH gene.  The invention also relates to methods for screening
CC   for HH homozygotes, to HH diagnosis, prenatal screening and
CC   diagnosis, and therapies of HH disease, including gene therapy,
CC   protein- and antibody-based therapeutics, and small molecule
CC   therapeutics.
XX
SQ   Sequence 1440 BP; 347 A; 355 C; 407 G; 331 T; 0 other;

  Query Match             48.6%;  Score 1219;  DB 18;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-295;
  Matches 1219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy    1201 CCTCAAAAAGATTTCCCCA 1219
           ||||||||||||||||||
Db    1422 CCTCAAAAAGATTTCCCCA 1440
```

RESULT 4
AAC68429

```
ID   AAC68429 standard; DNA; 1440 BP.
XX
AC   AAC68429;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis cDNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ss.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
```

```
DR   WPI; 2001-006341/01.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 1440 BP; 347 A; 355 C; 407 G; 331 T; 0 other;
```

```
  Query Match             48.6%;  Score 1219;  DB 22;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-295;
  Matches 1219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001
```

```
Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy    1201 CCTCAAAAAGATTTCCCCA 1219
           ||||||||||||||||||
Db    1422 CCTCAAAAAGATTTCCCCA 1440
```

RESULT 5
AAC68430
ID AAC68430 standard; DNA; 1440 BP.
XX
AC AAC68430;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1 mutation cDNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ss.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 4; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 1440 BP; 348 A; 355 C; 406 G; 331 T; 0 other;

Query Match 48.6%; Score 1217.4; DB 22; Length 1440;
Best Local Similarity 99.9%; Pred. No. 3.2e-295;
Matches 1218; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
Db   1422 CCTCAAAAAGATTTCCCCA 1440


RESULT 6
AAC68431
ID   AAC68431 standard; DNA; 1440 BP.
XX
AC   AAC68431;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d2 mutation cDNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ss.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 1440 BP; 347 A; 354 C; 408 G; 331 T; 0 other;

  Query Match             48.6%;  Score 1217.4;  DB 22;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 3.2e-295;
  Matches 1218;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301
```

```
Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
Db   1422 CCTCAAAAAGATTTCCCCA 1440


RESULT 7
AAC68432
ID   AAC68432 standard; DNA; 1440 BP.
XX
AC   AAC68432;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d1/2 mutation cDNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ss.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 4; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 1440 BP; 348 A; 354 C; 407 G; 331 T; 0 other;

  Query Match           48.5%;  Score 1215.8;  DB 22;  Length 1440;
  Best Local Similarity  99.8%;  Pred. No. 8.1e-295;
  Matches 1217;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
```

```
Db   1422 CCTCAAAAAGATTTCCCCA 1440


RESULT 8
AAL36747
ID   AAL36747 standard; DNA; 5749 BP.
XX
AC   AAL36747;
XX
DT   08-JAN-2002  (first entry)
XX
DE   Human musculoskeletal system related polynucleotide SEQ ID NO 3112.
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein;
KW   musculoskeletal system; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155367-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01338.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
```

```
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-451937/48.
XX
PT   Isolated polypeptide for treating, preventing and/ or prognosing
PT   disorders related to the musculoskeletal system including
PT   musculoskeletal cancers and also for testing and detection e.g.
PT   diagnosis -
XX
PS   Example 2; SEQ ID NO 3112; 781pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC   (ABB03087-ABB04109) associated with the musculoskeletal system useful
CC   for preventing, treating or ameliorating medical conditions e.g. by
CC   protein or gene therapy. The genes are isolated from a range of human
CC   tissues disclosed in the specification. The nucleic acids, proteins,
CC   antibodies and (ant)agonists are useful in the diagnosis, treatment
CC   and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC   other cancers of the adrenal gland, bone, bone marrow, breast,
CC   gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC   disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC   anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC   multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC   (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC   healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC   and (f) infectious diseases such as viral, bacterial, fungal and
CC   parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5749 BP; 1600 A; 1192 C; 1403 G; 1553 T; 1 other;

  Query Match          42.0%;  Score 1051.6;  DB 22;  Length 5749;
  Best Local Similarity  98.7%;  Pred. No. 2.3e-253;
  Matches 1060;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3103 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 3162

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3163 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 3222

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3223 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 3282

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3283 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 3342

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3343 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 3402

Qy   1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3403 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 3462

Qy   1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3463 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 3522

Qy   1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3523 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 3582

Qy   1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3583 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 3642

Qy   1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3643 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 3702

Qy   1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   3703 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 3762

Qy   1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3763 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 3822

Qy   1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3823 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 3882

Qy   1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3883 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 3942

Qy   1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3943 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 4002

Qy   1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4003 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 4062

Qy   1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4063 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 4122

Qy   2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
           |||||||||||||||||||||| ||||||| ||||||     || ||   ||
Db   4123 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 4176


RESULT 9
AAT96690
ID   AAT96690 standard; DNA; 10825 BP.
XX
AC   AAT96690;
XX
DT   14-APR-1998  (first entry)
XX
DE   Hereditary haemochromatosis gene.
XX
KW   Hereditary haemochromatosis; metal toxicity; diagnosis;
KW   gene therapy; prenatal screening; human; ds.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   CDS             361..7147
FT                   /*tag= a
FT                   /note= "contains introns"
FT   intron          437..3761
FT                   /*tag= b
FT                   /number= 1
FT   intron          4026..4234
FT                   /*tag= c
FT                   /number= 2
FT   intron          4511..5605
FT                   /*tag= d
FT                   /number= 3
FT   intron          5882..6039
FT                   /*tag= e
FT                   /number= 4
FT   intron          6154..7106
FT                   /*tag= f
FT                   /number= 5
FT   mutation        3872
FT                   /*tag= g
FT                   /note= "C to G substitution (24d2 mutation)
FT                         results in His to Asp substitution"
FT   variation       3878
FT                   /*tag= h
FT                   /note= "A to T substitution (24d7 variant)
FT                         results in Ser to Cys substitution"
FT   mutation        5834
FT                   /*tag= i
FT                   /note= "G to A substitution (24d1 mutation
FT                         associated with HH), results in Cys to
FT                         Tyr substitution"
XX
PN   WO9738137-A1.
XX
PD   16-OCT-1997.
XX
PF   04-APR-1997;   97WO-US06254.
XX
PR   23-MAY-1996;   96US-0652265.
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
XX
PA   (MERC-) MERCATOR GENETICS INC.
XX
PI   Drayna DT,  Feder JN,  Gnirke A,  Ruddy D,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1997-512743/47.
DR   P-PSDB; AAW36499.
XX
PT   Hereditary haemochromatosis gene and variants - useful for diagnosis
PT   and treatment of hereditary haemochromatosis disease
XX
PS   Disclosure; Fig 3; 115pp; English.
XX
CC   This genomic DNA sequence corresponds to the human gene whose
CC   mutated form is associated with hereditary haemochromatosis (HH).
CC   To identify this novel gene, allelic association patterns were
CC   determined between known markers and the HH locus in the HLA region
CC   of chromosome 6.  A physical clone coverage was then generated
CC   extending from D6S265, which is a marker that is centromeric of
CC   HLA-A, in a telomeric direction through D6S276, a marker at which
CC   the allelic association was no longer observed.  A single mutation
CC   (24d1) in the HH gene appears responsible for the majority of HH
CC   disease.  This comprises a G to A substitution that is present in
CC   86% of affected chromosomes and in 4% of unaffected chromosomes.
CC   It results in a Cys to Tyr substitution in the encoded protein (see
CC   AAW36499) at a critical disulphide bridge important for secondary
CC   structure.  The following are claimed: the HH genomic DNA (I), a
CC   1437 bp cDNA sequence (Ia) (see AAT96691) and their 24d1, 24d2 and
CC   24d7 variants; a cloning or expression vector; host cells; a
CC   peptide product chosen from the HH gene product, its variants
CC   (24d1, 24d2 and 24d7), or a peptide of at least 56 amino acid
CC   residues of these; an antibody produced using the peptide; a method
CC   to determine the presence or absence of the common HH gene
CC   mutation; an animal model for the HH disease; metal chelation
CC   agents, T-cell differentiation factors and therapeutic agents for
CC   the mitigation of injury due to oxidative process in vivo or
CC   mitigation of iron overload; a method for screening potential
CC   therapeutic agents for activity in connection with HH disease; an
CC   antisense oligonucleotide directed against a transcriptional
CC   product of a nucleic acid sequence as above; and oligonucleotides
CC   or pairs of oligonucleotides covering a range of nucleotides from
CC   (I), (Ia) or their variants, useful for detecting a polymorphism in
CC   the HH gene.  The invention also relates to methods for screening
CC   for HH homozygotes, to HH diagnosis, prenatal screening and
CC   diagnosis, and therapies of HH disease, including gene therapy,
CC   protein- and antibody-based therapeutics, and small molecule
CC   therapeutics.
XX
SQ   Sequence 10825 BP; 2996 A; 2254 C; 2648 G; 2927 T; 0 other;

  Query Match             42.0%;  Score 1051.6;  DB 18;  Length 10825;
  Best Local Similarity   98.7%;  Pred. No. 3.1e-253;
  Matches 1060;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7104 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 7163

Qy    1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7164 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 7223

Qy    1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7224 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 7283

Qy    1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7284 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 7343

Qy    1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7344 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 7403

Qy    1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7404 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 7463

Qy    1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7464 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 7523

Qy    1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7524 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 7583

Qy    1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7584 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 7643

Qy    1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7644 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 7703

Qy    1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    7704 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 7763

Qy    1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7764 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 7823

Qy    1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7824 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 7883

Qy   1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7884 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 7943

Qy   1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7944 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 8003

Qy   1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8004 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 8063

Qy   1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8064 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 8123

Qy   2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
          ||||||||||||||||||||||| |||||||  ||||||     || ||   ||
Db   8124 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 8177
```

RESULT 10
AAC68425
ID AAC68425 standard; DNA; 10825 BP.
XX
AC AAC68425;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis DNA.
XX
KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36869.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2998 A; 2253 C; 2648 G; 2926 T; 0 other;

Query Match 42.0%; Score 1051.6; DB 22; Length 10825;

Best Local Similarity 98.7%; Pred. No. 3.1e-253;
Matches 1060; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
Qy   1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7104 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 7163

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7164 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 7223

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7224 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 7283

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7284 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 7343

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7344 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 7403

Qy   1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7404 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 7463

Qy   1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7464 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 7523

Qy   1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7524 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 7583

Qy   1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7584 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 7643

Qy   1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7644 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 7703

Qy   1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   7704 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 7763

Qy   1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7764 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 7823

Qy   1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7824 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 7883

Qy   1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7884 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 7943

Qy   1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7944 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 8003

Qy   1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8004 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 8063

Qy   1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8064 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 8123

Qy   2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
          ||||||||||||||||||||||| |||||||  ||||||     || ||   ||
```

```
Db   8124 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 8177


RESULT 11
AAC68426
ID   AAC68426 standard; DNA; 10825 BP.
XX
AC   AAC68426;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d1 mutation DNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36870.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 10825 BP; 2999 A; 2253 C; 2647 G; 2926 T; 0 other;

  Query Match             42.0%;  Score 1051.6;  DB 22;  Length 10825;
  Best Local Similarity   98.7%;  Pred. No. 3.1e-253;
  Matches 1060;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7104 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 7163

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7164 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 7223

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7224 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 7283

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7284 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 7343

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7344 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 7403

Qy   1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7404 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 7463

Qy   1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7464 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 7523

Qy   1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7524 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 7583

Qy   1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7584 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 7643

Qy   1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7644 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 7703

Qy   1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   7704 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 7763

Qy   1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7764 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 7823

Qy   1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7824 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 7883

Qy   1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7884 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 7943

Qy   1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7944 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 8003

Qy   1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8004 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 8063

Qy   1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8064 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 8123

Qy   2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
          ||||||||||||||||||||| ||||||| ||||||    || ||   ||
Db   8124 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 8177


RESULT 12
AAC68427
ID   AAC68427 standard; DNA; 10825 BP.
XX
AC   AAC68427;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d2 mutation DNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
```

```
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36871.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 10825 BP; 2998 A; 2252 C; 2649 G; 2926 T; 0 other;

  Query Match             42.0%;  Score 1051.6;  DB 22;  Length 10825;
  Best Local Similarity   98.7%;  Pred. No. 3.1e-253;
  Matches 1060;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7104 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 7163

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7164 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 7223

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7224 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 7283

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7284 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 7343

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7344 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 7403

Qy   1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7404 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 7463

Qy   1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7464 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 7523

Qy   1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7524 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 7583

Qy   1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7584 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 7643

Qy   1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7644 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 7703

Qy   1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   7704 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 7763

Qy   1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7764 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 7823

Qy   1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7824 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 7883

Qy   1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7884 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 7943

Qy   1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7944 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 8003

Qy   1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8004 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 8063

Qy   1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8064 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 8123

Qy   2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
           |||||||||||||||||||| |||||||  ||||||     || ||   ||
Db   8124 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 8177


RESULT 13
AAC68428
ID   AAC68428 standard; DNA; 10825 BP.
XX
AC   AAC68428;
XX
DT   21-FEB-2001  (first entry)
XX
DE   Human hereditary hemochromatosis 24d1/2 mutation DNA.
XX
KW   HH; hereditary hemochromatosis; chelation agent;
KW   T-cell differentiation factor; iron overload; ds.
XX
OS   Homo sapiens.
XX
PN   US6140305-A.
XX
PD   31-OCT-2000.
XX
PF   04-APR-1997;   97US-0834497.
XX
PR   04-APR-1996;   96US-0630912.
PR   16-APR-1996;   96US-0632673.
PR   23-MAY-1996;   96US-0652265.
XX
PA   (BIRA ) BIO-RAD LAB INC.
XX
PI   Thomas WJ,  Drayna DT,  Gnirke A,  Ruddy D,  Tsuchihashi Z,  Wolff RK;
PI   Feder JN;
XX
DR   WPI; 2001-006341/01.
DR   P-PSDB; AAB36872.
XX
PT   New hereditary hemochromatosis gene products or polypeptides, useful
PT   for treating hereditary hemochromatosis in a patient, and as a metal
PT   chelation agent alleviating iron overload -
XX
PS   Disclosure; Fig 3; 108pp; English.
XX
```

```
CC   The present invention relates to hereditary hemochromatosis gene
CC   products. These proteins may be used to treat a patient diagnosed as
CC   having human hemochromatosis disease. It is also useful as a metal
CC   chelation agent or as a T-cell differentiation factor, and for
CC   alleviating iron overload. They may also be used in protein replacement
CC   therapy for individuals having a defective human hemochromatosis gene.
XX
SQ   Sequence 10825 BP; 2999 A; 2252 C; 2648 G; 2926 T; 0 other;

  Query Match            42.0%;  Score 1051.6;  DB 22;  Length 10825;
  Best Local Similarity  98.7%;  Pred. No. 3.1e-253;
  Matches 1060;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7104 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 7163

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7164 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 7223

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7224 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 7283

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7284 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 7343

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7344 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 7403

Qy   1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7404 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 7463

Qy   1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7464 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 7523

Qy   1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7524 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 7583

Qy   1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7584 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 7643

Qy   1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7644 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 7703

Qy   1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   7704 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 7763

Qy   1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7764 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 7823

Qy   1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7824 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 7883

Qy   1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7884 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 7943

Qy   1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7944 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 8003

Qy   1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8004 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 8063

Qy   1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8064 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 8123

Qy   2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
           ||||||||||||||||||||||| |||||||  ||||||     || ||   ||
Db   8124 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 8177


RESULT 14
AAA96794
ID   AAA96794 standard; cDNA; 12146 BP.
XX
AC   AAA96794;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Genomic DNA of a histocompatibility iron loading (HFE) gene.
XX
KW   Human; histocompatibility iron loading protein; HFE protein;
KW   major histocompatibility complex; non-classical class I gene;
KW   chromosome 6p; iron disorder; haemochromatosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            1028..1324
FT                   /*tag= a
FT                   /number= 1
FT   intron          1325..4651
FT                   /*tag= b
FT                   /number= 1
FT   exon            4652..4915
FT                   /*tag= c
FT                   /number= 2
FT   intron          4916..5124
FT                   /*tag= d
FT                   /number= 2
FT   exon            5125..5400
FT                   /*tag= e
FT                   /number= 3
FT   intron          5401..6493
FT                   /*tag= f
FT                   /number= 3
FT   exon            6494..6769
FT                   /*tag= g
FT                   /number= 4
FT   intron          6770..6927
FT                   /*tag= h
FT                   /number= 4
FT   exon            6928..7041
FT                   /*tag= i
FT                   /number= 5
FT   intron          7042..7994
FT                   /*tag= j
FT                   /number= 5
FT   exon            7995..9050
FT                   /*tag= k
FT                   /number= 6
FT   intron          9051..10205
FT                   /*tag= l
FT                   /number= 6
FT   exon            10206..10637
FT                   /*tag= m
XX
PN   WO200058515-A1.
XX
PD   05-OCT-2000.
XX
```

```
PF   24-MAR-2000; 2000WO-US07982.
XX
PR   26-MAR-1999;   99US-0277457.
XX
PA   (BILL-) BILLUPS-ROTHENBERG INC.
XX
PI   Rothenberg BE,  Sawada-Hirai R,  Barton JC;
XX
DR   WPI; 2000-647244/62.
XX
PT   Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT   susceptibility to develop it, by determining the presence of a mutation
PT   in exon 2 or an intron of a histocompatibility iron loading nucleic
PT   acid -
XX
PS   Example 1; Page 21-28; 55pp; English.
XX
CC   The present sequence represents the human histocompatibility iron
CC   loading (HFE) gene. The HFE gene is a major histocompatibility (MHC)
CC   non-classical class I gene located on chromosome 6p. Mutations in the
CC   gene lead to iron disorders. The specification describes a method for
CC   diagnosing an iron disorder or a genetic susceptibility to develop the
CC   disorder in a mammal. The method comprises determining the presence of
CC   a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC   is not a C to G missense mutation at nucleotide 187 of the sequence
CC   given in A96769 (Genbank Accession number U60319). The presence of the
CC   mutation indicates the disorder or the genetic susceptibility to the
CC   disorder. The method is used to diagnose an iron disorder
CC   e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
SQ   Sequence 12146 BP; 3383 A; 2474 C; 2911 G; 3378 T; 0 other;

  Query Match            42.0%;  Score 1051.6;  DB 21;  Length 12146;
  Best Local Similarity  98.7%;  Pred. No. 3.3e-253;
  Matches 1060;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7992 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 8051

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8052 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 8111

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8112 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 8171

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8172 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 8231

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8232 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 8291

Qy   1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8292 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 8351

Qy   1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8352 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 8411

Qy   1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8412 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 8471

Qy   1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8472 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 8531

Qy   1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8532 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 8591

Qy   1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   8592 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 8651

Qy   1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8652 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 8711

Qy   1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8712 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 8771

Qy   1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8772 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 8831

Qy   1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8832 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 8891

Qy   1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8892 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 8951

Qy   1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8952 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 9011

Qy   2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
           ||||||||||||||||||||||| ||||||| ||||||     || ||   ||
Db   9012 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 9065


RESULT 15
AAV57903/c
ID   AAV57903 standard; DNA; 237326 BP.
XX
AC   AAV57903;
XX
DT   21-DEC-1998  (first entry)
XX
DE   Hereditary haemochromatosis subregion from an HH affected individual.
XX
KW   Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
KW   diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;
KW   BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
KW   type 1 sodium transport gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO9814466-A1.
XX
PD   09-APR-1998.
XX
PF   30-SEP-1997;   97WO-US17658.
XX
PR   07-MAY-1997;   97US-0852495.
PR   01-OCT-1996;   96US-0724394.
XX
PA   (PROG-) PROGENTIOR INC.
XX
PI   Feder JN,  Kronmal GS,  Lauer PM,  Ruddy DA,  Thomas WJ;
PI   Tsuchihashi Z,  Wolff RK;
XX
DR   WPI; 1998-240014/21.
XX
PT   Hereditary haemochromatosis gene products - used to develop products
PT   for the diagnosis and treatment of hereditary disorders in iron
PT   metabolism
XX
```

PS Claim 1; Fig 9; 209pp; English.
XX
CC The present invention describes hereditary haemochromatosis gene
CC products from the human haemochromatosis gene. The present sequence
CC represents a hereditary haemochromatosis subregion from an hereditary
CC haemochromatosis (HH) affected individual. Also described is a
CC method to determine the presence or absence of the common hereditary
CC haemochromatosis (HFE) gene mutation in an individual comprising:
CC (a) providing DNA or RNA from the individual; and (b) assessing the
CC DNA or RNA for the presence or absence of a haplotype or genotype where
CC the presence or absence of the haplotype genotype indicates the likely
CC presence of the HFE gene mutation in the genome of the individual. The
CC HFE gene sequences from the present invention can be used to develop
CC products for use in the diagnosis and treatment of HFE. The present
CC invention also describes BTF genes, which are homologues of the milk
CC protein butyrophilin (BT), and can be used in the production of agonists
CC and antagonists of BT function. Also described are: (1) a RoRet gene
CC which can be used to develop products for the study, diagnosis and
CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
CC which are homologues of a type 1 sodium transport gene, and can
CC similarly be used for hypophosphatemia.
XX
SQ Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;

Query Match 41.8%; Score 1048.4; DB 19; Length 237326;
Best Local Similarity 98.5%; Pred. No. 7.8e-252;
Matches 1058; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy   1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39998 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 39939

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39938 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 39879

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39878 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 39819

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39818 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 39759

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39758 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 39699

Qy   1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39698 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 39639

Qy   1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39638 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 39579

Qy   1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39578 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 39519

Qy   1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39518 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 39459

Qy   1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39458 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 39399

Qy   1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
          |||||||||||||||| |||||||||||||||||||| ||||||||||||||||||||
Db  39398 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGGACCTGTCCCAGAAAAAGCATC 39339

Qy   1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  39338 ATGGCTATCTCTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 39279

Qy   1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39278 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 39219

Qy   1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39218 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 39159

Qy   1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39158 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 39099

Qy   1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39098 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 39039

Qy   1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39038 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 38979

Qy   2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
          |||||||||||||||||||||| ||||||| ||||||     || ||    ||
Db  38978 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 38925
```

Search completed: November 2, 2002, 04:19:13
Job time : 1014.45 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:55:33 ; Search time 99.3352 Seconds
(without alignments)
6196.774 Million cell updates/sec

Title: US-09-981-606-1
Perfect score: 2506
Sequence: 1 atgggcccgcgagccaggcc.....ttgtattgtaaaaaaaaaaa 2506

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2506	100.0	2506	4	US-09-277-457-1	Sequence 1, Appli
2	1219	48.6	1440	3	US-08-652-265-9	Sequence 9, Appli
3	1219	48.6	1440	3	US-08-834-497A-9	Sequence 9, Appli
4	1219	48.6	1440	4	US-09-503-444A-9	Sequence 9, Appli
5	1217.4	48.6	1440	3	US-08-652-265-10	Sequence 10, Appl
6	1217.4	48.6	1440	3	US-08-652-265-11	Sequence 11, Appl
7	1217.4	48.6	1440	3	US-08-834-497A-10	Sequence 10, Appl
8	1217.4	48.6	1440	3	US-08-834-497A-11	Sequence 11, Appl
9	1217.4	48.6	1440	4	US-09-503-444A-10	Sequence 10, Appl
10	1217.4	48.6	1440	4	US-09-503-444A-11	Sequence 11, Appl
11	1215.8	48.5	1440	3	US-08-652-265-12	Sequence 12, Appl
12	1215.8	48.5	1440	3	US-08-834-497A-12	Sequence 12, Appl
13	1215.8	48.5	1440	4	US-09-503-444A-12	Sequence 12, Appl
14	1051.6	42.0	10825	3	US-08-652-265-1	Sequence 1, Appli
15	1051.6	42.0	10825	3	US-08-652-265-3	Sequence 3, Appli
16	1051.6	42.0	10825	3	US-08-652-265-5	Sequence 5, Appli
17	1051.6	42.0	10825	3	US-08-652-265-7	Sequence 7, Appli
18	1051.6	42.0	10825	3	US-08-834-497A-1	Sequence 1, Appli
19	1051.6	42.0	10825	3	US-08-834-497A-3	Sequence 3, Appli
20	1051.6	42.0	10825	3	US-08-834-497A-5	Sequence 5, Appli
21	1051.6	42.0	10825	3	US-08-834-497A-7	Sequence 7, Appli
22	1051.6	42.0	10825	4	US-09-503-444A-1	Sequence 1, Appli
23	1051.6	42.0	10825	4	US-09-503-444A-3	Sequence 3, Appli
24	1051.6	42.0	10825	4	US-09-503-444A-5	Sequence 5, Appli
25	1051.6	42.0	10825	4	US-09-503-444A-7	Sequence 7, Appli
26	1051.6	42.0	12146	4	US-09-277-457-27	Sequence 27, Appl
27	1051.6	42.0	246240	2	US-08-724-394A-20	Sequence 20, Appl
28	1051.6	42.0	246240	2	US-08-724-394A-21	Sequence 21, Appl
29	1051.6	42.0	246240	2	US-08-724-394A-22	Sequence 22, Appl
30	280	11.2	360	3	US-08-905-124-5	Sequence 5, Appli
31	280	11.2	517	1	US-08-632-673B-3	Sequence 3, Appli
32	280	11.2	517	3	US-08-652-265-20	Sequence 20, Appl
33	280	11.2	517	3	US-08-834-497A-20	Sequence 20, Appl
34	280	11.2	517	4	US-09-503-444A-20	Sequence 20, Appl
35	279	11.1	517	1	US-08-632-673B-13	Sequence 13, Appl
36	278.4	11.1	517	1	US-08-632-673B-4	Sequence 4, Appli
37	278.4	11.1	517	3	US-08-652-265-21	Sequence 21, Appl
38	278.4	11.1	517	3	US-08-834-497A-21	Sequence 21, Appl
39	278.4	11.1	517	4	US-09-503-444A-21	Sequence 21, Appl
40	175.2	7.0	1112	3	US-08-890-719-5	Sequence 5, Appli
41	173.6	6.9	1230	3	US-08-890-719-9	Sequence 9, Appli
42	170.4	6.8	1195	3	US-08-890-719-10	Sequence 10, Appl
43	168.4	6.7	1197	3	US-08-890-719-7	Sequence 7, Appli
44	168.4	6.7	1262	3	US-08-890-719-37	Sequence 37, Appl
45	163.6	6.5	1145	3	US-08-890-719-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-277-457-1
; Sequence 1, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
; APPLICANT: Rothenberg, Barry E.
; APPLICANT: Sawada-Hirai, Ritsuko
; APPLICANT: Barton, James C.
; TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
; FILE REFERENCE: 10653/002001
; CURRENT APPLICATION NUMBER: US/09/277,457
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2506
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (0)...(0)
; OTHER INFORMATION: Missense mutation at nucleotide 314
US-09-277-457-1

Query Match 100.0%; Score 2506; DB 4; Length 2506;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200

Qy   1201 CCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTG 1260

Qy   1261 TGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCA 1320

Qy   1321 TTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAG 1380

Qy   1381 GACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAAGCTTTGAAC 1440

Qy   1441 CCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACATGTATCTATGCATTTTCT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACATGTATCTATGCATTTTCT 1500

Qy   1501 GGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCA 1560

Qy   1561 CCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCACAGCTATGAAGGCTGT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGCACAGCTATGAAGGCTGT 1620

Qy   1621 GCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTGGGTA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATCATGGCTATCTGTGGGTA 1680

Qy   1681 GTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGT 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAAAGGGGTTGTGAAGAGGT 1740

Qy   1741 GTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCAT 1800

Qy   1801 TTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATAATTTTCTACCTGGTCTC 1860

Qy   1861 TCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGCACTTACTTCGTGTCCGA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGCACTTACTTCGTGTCCGA 1920

Qy   1921 CTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTTACAATAATTCTATGAG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCTTACAATAATTCTATGAG 1980

Qy   1981 ATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGCACGG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGCACGG 2040

Qy   2041 TGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCC 2100

Qy   2101 GTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATG 2160

Qy   2161 CACTTGGCTGCATAAATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATA 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 CACTTGGCTGCATAAATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATA 2220

Qy   2221 CCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCT 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCT 2280

Qy   2281 CTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATG 2340

Qy   2341 TGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATT 2400

Qy   2401 AGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 AGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAA 2460

Qy   2461 AATGCATATACTTTAATAAATGTACATTGTATTGTAAAAAAAAAAA 2506
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AATGCATATACTTTAATAAATGTACATTGTATTGTAAAAAAAAAAA 2506
```

```
RESULT 2
US-08-652-265-9
; Sequence 9, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "c")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(414, "a")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d7
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "g")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-9

  Query Match             48.6%;  Score 1219;  DB 3;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
                                                                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy    1201 CCTCAAAAAGATTTCCCCA 1219
           ||||||||||||||||||
Db    1422 CCTCAAAAAGATTTCCCCA 1440


RESULT 3
US-08-834-497A-9
; Sequence 9, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "c")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(414, "a")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d7
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "g")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-08-834-497A-9

  Query Match             48.6%;  Score 1219;  DB 3;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001
```

```
Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy    1201 CCTCAAAAAGATTTCCCCA 1219
           ||||||||||||||||||
Db    1422 CCTCAAAAAGATTTCCCCA 1440
```

```
RESULT 4
US-09-503-444A-9
; Sequence 9, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "c")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(414, "a")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d7
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "g")
;      OTHER INFORMATION:  /phenotype= "normal or wild-type
;      OTHER INFORMATION:  (unaffected)"
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-9

  Query Match             48.6%;  Score 1219;  DB 4;  Length 1440;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701
```

```
Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
Db   1422 CCTCAAAAAGATTTCCCCA 1440
```

RESULT 5
US-08-652-265-10
; Sequence 10, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(1066, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d1
US-08-652-265-10

Query Match 48.6%; Score 1217.4; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
                |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
                ||||||||||||||||||
Db   1422 CCTCAAAAAGATTTCCCCA 1440
```

```
RESULT 6
US-08-652-265-11
; Sequence 11, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-08-652-265-11

  Query Match             48.6%;  Score 1217.4;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1218;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761
```

```
Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
Db   1422 CCTCAAAAAGATTTCCCCA 1440
```

```
RESULT 7
US-08-834-497A-10
; Sequence 10, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-834-497A-10

  Query Match             48.6%;  Score 1217.4;  DB 3;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1218;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy      1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy     61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581
```

```
Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy    1201 CCTCAAAAAGATTTCCCCA 1219
           ||||||||||||||||||
Db    1422 CCTCAAAAAGATTTCCCCA 1440
```

```
RESULT 8
US-08-834-497A-11
; Sequence 11, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-08-834-497A-11

  Query Match           48.6%;  Score 1217.4;  DB 3;  Length 1440;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 1218;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
```

```
Db    402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
Db   1422 CCTCAAAAAGATTTCCCCA 1440
```

RESULT 9

```
US-09-503-444A-10
; Sequence 10, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-10

  Query Match             48.6%;  Score 1217.4;  DB 4;  Length 1440;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1218;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
```

```
Db    282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy    121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
Db    342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy    181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
Db    402 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy    241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
Db    462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy    301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
Db    522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy    361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
Db    582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy    421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
Db    642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy    481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
Db    702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy    541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
Db    762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy    601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
Db    822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy    661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
Db    882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy    721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
Db    942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy    781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
Db   1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy    841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
Db   1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy    901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
Db   1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy    961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
Db   1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy   1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
Db   1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy   1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
Db   1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy   1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
Db   1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy   1201 CCTCAAAAAGATTTCCCCA 1219
Db   1422 CCTCAAAAAGATTTCCCCA 1440
```

```
RESULT 10
US-09-503-444A-11
; Sequence 11, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
US-09-503-444A-11

  Query Match             48.6%;  Score 1217.4;  DB 4;  Length 1440;
```

```
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1218;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTG 1181

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy    1201 CCTCAAAAAGATTTCCCCA 1219
           ||||||||||||||||||
Db    1422 CCTCAAAAAGATTTCCCCA 1440


RESULT 11
US-08-652-265-12
; Sequence 12, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
```

```
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-12

  Query Match             48.5%;  Score 1215.8;  DB 3;  Length 1440;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1217;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTG 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTTGTCGTCATCTTG 1181

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy    1201 CCTCAAAAAGATTTCCCCA 1219
          ||||||||||||||||||
Db    1422 CCTCAAAAAGATTTCCCCA 1440


RESULT 12
US-08-834-497A-12
; Sequence 12, Application US/08834497A
; Patent No. 6140305
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/834,497A
;      FILING DATE:  04-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/632,673
;      FILING DATE:  16-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/630,912
;      FILING DATE:  04-APR-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0056-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
```

```
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  222..1268
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(408, "g")
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(1066, "a")
;       OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /label= 24d1
US-08-834-497A-12

  Query Match             48.5%;  Score 1215.8;  DB 3;  Length 1440;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1217;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy    1201 CCTCAAAAAGATTTCCCCA 1219
           ||||||||||||||||||||
Db    1422 CCTCAAAAAGATTTCCCCA 1440


RESULT 13
US-09-503-444A-12
; Sequence 12, Application US/09503444A
; Patent No. 6228594
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect Version 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/503,444A
;      FILING DATE:  14-Feb-2000
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/652,265
;      FILING DATE:  23-May-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,673
;      FILING DATE:  16-Apr-1996
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  08/630,912
;      FILING DATE:  04-Apr-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0088-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-790-9090
;      TELEFAX:  212-869-9741
;      TELEX:  66141
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1440 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  222..1268
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(408, "g")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d2
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(1066, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-09-503-444A-12

  Query Match             48.5%;  Score 1215.8;  DB 4;  Length 1440;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1217;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     222 ATGGGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTG 281

Qy      61 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     282 CAGGGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAG 341

Qy     121 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     342 GACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTC 401

Qy     181 TATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 240
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     402 TATGATGATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461

Qy     241 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     462 AGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTT 521

Qy     301 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     522 GACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGTCCCACACCCTGCAGGTC 581

Qy     361 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     582 ATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTAT 641

Qy     421 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     642 GATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCC 701

Qy     481 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     702 AGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGAACAGG 761

Qy     541 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     762 GCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGT 821

Qy     601 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     822 GTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881

Qy     661 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     882 ACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAACATCACCATGAAGTGGCTG 941

Qy     721 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942 AAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001

Qy     781 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1002 GATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCTGGGGAAGAGCAGAGATAT 1061

Qy     841 ACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 900
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1062 ACGTACCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATTGTGATCTGGGAGCCCTCA 1121

Qy     901 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1122 CCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCTGTTTTTGTCGTCATCTTG 1181

Qy     961 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1182 TTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAGGGTTCAAGAGGAGCCATGGGG 1241

Qy    1021 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1242 CACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGACTCACTGTGGGAAGGAGA 1301

Qy    1081 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1302 CAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361

Qy    1141 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1362 TGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTT 1421

Qy    1201 CCTCAAAAAGATTTCCCCA 1219
           ||||||||||||||||||
Db    1422 CCTCAAAAAGATTTCCCCA 1440


RESULT 14
US-08-652-265-1
; Sequence 1, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/652,265
;       FILING DATE:  23-MAY-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Smith, William M.
;       REGISTRATION NUMBER:  30,223
;       REFERENCE/DOCKET NUMBER:  17957-000500
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 576-0200
;       TELEFAX:  (415) 576-0300
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10825 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;       LOCATION:  6040..6153, 7107..7147)
;       OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;       OTHER INFORMATION:
;       OTHER INFORMATION:  /note= "No. 6025130mal or wild-type (unaffected)
;       OTHER INFORMATION:  Hereditary Hemochromatosis (HH) gene
;       OTHER INFORMATION:  allele"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  140..7319
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) allele
;       OTHER INFORMATION:  cDNA (SEQ ID NO:9)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  3852..3891
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d2(C)
;       OTHER INFORMATION:  allele (SEQ ID NO:41)"
;     FEATURE:
;       NAME/KEY:  -
;       LOCATION:  5507..6023
;       OTHER INFORMATION:  /note= "start and stop positions for
;       OTHER INFORMATION:  normal or wild-type (unaffected) genomic
;       OTHER INFORMATION:  sequence surrounding variant for 24d1(G)
;       OTHER INFORMATION:  allele (SEQ ID NO:20)"
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3872, "c")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d2
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(3878, "a")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d7
;     FEATURE:
;       NAME/KEY:  allele
;       LOCATION:  replace(5834, "g")
;       OTHER INFORMATION:  /phenotype= "normal or wild-type
;       OTHER INFORMATION:  (unaffected)"
;       OTHER INFORMATION:  /label= 24d1
US-08-652-265-1

  Query Match           42.0%;  Score 1051.6;  DB 3;  Length 10825;
  Best Local Similarity   98.7%;  Pred. No. 5.4e-282;
  Matches 1060;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;
```

```
Qy   1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7104 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 7163

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7164 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 7223

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7224 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 7283

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7284 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 7343

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7344 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 7403

Qy   1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7404 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 7463

Qy   1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7464 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 7523

Qy   1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7524 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 7583

Qy   1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7584 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 7643

Qy   1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7644 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 7703

Qy   1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   7704 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 7763

Qy   1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7764 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 7823

Qy   1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7824 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 7883

Qy   1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7884 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 7943

Qy   1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7944 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 8003

Qy   1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8004 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 8063

Qy   1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8064 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 8123

Qy   2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
          |||||||||||||||||||||| ||||||  ||||||     || ||   ||
Db   8124 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 8177
```

```
RESULT 15
US-08-652-265-3
; Sequence 3, Application US/08652265
; Patent No. 6025130
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas, Winston J.
;    APPLICANT:  Drayna, Dennis T.
;    APPLICANT:  Feder, John N.
;    APPLICANT:  Gnirke, Andreas
;    APPLICANT:  Ruddy, David
;    APPLICANT:  Tsuchihashi, Zenta
;    APPLICANT:  Wolff, Roger K.
;    TITLE OF INVENTION:  Hereditary Hemochromatosis Gene
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,265
;      FILING DATE:  23-MAY-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  17957-000500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10825 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(361..436, 3762..4025, 4235..4510, 5606..5881,
;      LOCATION:  6040..6153, 7107..7147)
;      OTHER INFORMATION:  /product= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  mutation"
;      OTHER INFORMATION:  /note= "Hereditary Hemochromatosis (HH)
;      OTHER INFORMATION:  gene 24d1 allele"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  140..7319
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  24d1 allele cDNA (SEQ ID NO:10)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  3852..3891
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d2(C) allele (SEQ ID NO:41)"
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  5507..6023
;      OTHER INFORMATION:  /note= "start and stop positions for
;      OTHER INFORMATION:  genomic sequence surrounding variant
;      OTHER INFORMATION:  for 24d1(A) allele (SEQ ID NO:21)"
;    FEATURE:
;      NAME/KEY:  allele
;      LOCATION:  replace(5834, "a")
;      OTHER INFORMATION:  /phenotype= "Hereditary Hemochromatosis
;      OTHER INFORMATION:
;      OTHER INFORMATION:  /label= 24d1
US-08-652-265-3

  Query Match             42.0%;  Score 1051.6;  DB 3;  Length 10825;
  Best Local Similarity   98.7%;  Pred. No. 5.4e-282;
  Matches 1060;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1004 CAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 1063
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7104 CAGGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGCAGAC 7163

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7164 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 7223

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7224 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 7283

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7284 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 7343

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7344 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 7403

Qy   1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7404 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 7463

Qy   1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7464 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 7523

Qy   1424 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 1483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7524 TGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAACCTTACCAGATTTTTACACAT 7583

Qy   1484 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 1543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7584 GTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCTCTCTGTGTTACC 7643

Qy   1544 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 1603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7644 CAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGTGATGTGAGTTGC 7703

Qy   1604 ACAGCTATGAAGGCTGTGCACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 1663
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   7704 ACAGCTATGAAGGCTGTACACTGCACGAATGGAAGAGGCACCTGTCCCAGAAAAAGCATC 7763

Qy   1664 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 1723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7764 ATGGCTATCTGTGGGTAGTATGATGGGTGTTTTTAGCAGGTAGGAGGCAAATATCTTGAA 7823

Qy   1724 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 1783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7824 AGGGGTTGTGAAGAGGTGTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAG 7883

Qy   1784 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 1843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7884 TTTAATGGTGCCTTCATTTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAATA 7943

Qy   1844 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 1903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7944 ATTTTCTACCTGGTCTCTCCTTGTTCTGATAATGAAAATTATGATAAGGATGATAAAAGC 8003

Qy   1904 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 1963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8004 ACTTACTTCGTGTCCGACTCTTCTGAGCACCTACTTACATGCATTACTGCATGCACTTCT 8063
```

```
Qy    1964 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 2023
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    8064 TACAATAATTCTATGAGATAGGTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTG 8123

Qy    2024 AAGTAGGCCGGGCACGGTGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATTGCA 2077
           |||||||||||||||||||||| |||||||  ||||||      || ||    ||
Db    8124 AAGTAGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCA 8177
```

Search completed: November 2, 2002, 06:44:29
Job time : 140.335 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: November 2, 2002, 03:10:18 ; Search time 3791.46 Seconds
(without alignments)
8920.945 Million cell updates/sec

Title: US-09-981-606-1
Perfect score: 2506
Sequence: 1 atgggcccgcgagccaggcc.....ttgtattgtaaaaaaaaaaa 2506

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		738.6	29.5	819	10	BG747345	BG747345 602704818
2		640.8	25.6	1723	11	AK009581	AK009581 Mus muscu
3		506.8	20.2	570	10	BE272926	BE272926 601171213
c 4		451	18.0	520	10	N93736	N93736 zb51g06.s1
c 5		437.8	17.5	520	9	AI949947	AI949947 wq04f05.x
c 6		435.8	17.4	440	9	AW469921	AW469921 ha27e08.x
c 7		434.2	17.3	439	9	AA876054	AA876054 nr14d03.s
c 8		433.8	17.3	438	9	AI127651	AI127651 qc30h07.x
c 9		388.8	15.5	392	9	AI122894	AI122894 ok31a01.s
10		380.2	15.2	805	10	BG402460	BG402460 602466163
c 11		375.2	15.0	380	9	AI040303	AI040303 oy09c09.x
c 12		365.8	14.6	393	9	AI763178	AI763178 wi65f02.x
c 13		358	14.3	384	10	BF883952	BF883952 PM4-ET020
14		341.2	13.6	368	10	BF739151	BF739151 CM4-KT003
c 15		333.8	13.3	388	10	BF445847	BF445847 7p38b08.x
c 16		324.2	12.9	334	9	AW902003	AW902003 QV0-NN102
c 17		316.4	12.6	318	10	BF446089	BF446089 7p32a03.x
18		281.4	11.2	308	9	AA319758	AA319758 EST22021
19		277.8	11.1	550	10	BI339179	BI339179 364041 MA
c 20		274.2	10.9	299	10	R07647	R07647 ye98g01.s1
21		259.4	10.4	489	10	BE994943	BE994943 UI-M-CG0p
22		251.4	10.0	457	9	AI850020	AI850020 UI-M-BG0-
23		248.4	9.9	455	10	BE995172	BE995172 UI-M-CG0p
24		243.8	9.7	419	10	R07696	R07696 ye98g01.r1
25		239.2	9.5	343	10	R47761	R47761 yj62c07.r1
26		238.8	9.5	523	10	BF080089	BF080089 230846 MA
27		236.2	9.4	538	9	BE232472	BE232472 137492 MA
c 28		231.6	9.2	464	9	AA217236	AA217236 mu89b05.r
29		231.2	9.2	268	10	W21141	W21141 zb51g06.r1
30		217.6	8.7	831	10	BI452668	BI452668 603169877
31		214.6	8.6	546	10	BE809138	BE809138 214520 MA
c 32		202.6	8.1	258	10	R50398	R50398 yj62c07.s1
33		199.4	8.0	481	9	BB851691	BB851691 BB851691
34		198.2	7.9	392	10	BF465475	BF465475 UI-M-CG0p
35		187.4	7.5	1208	11	BC005306	BC005306 Homo sapi
36		186.8	7.5	502	9	BB858165	BB858165 BB858165
c 37		177.4	7.1	457	9	AI160732	AI160732 qb49b10.x
38		174	6.9	394	10	BF464345	BF464345 UI-M-CG0p
c 39		170.4	6.8	444	12	AZ025590	AZ025590 RPCI-23-3
c 40		156.8	6.3	536	12	AZ074871	AZ074871 RPCI-23-4
41		153.6	6.1	1016	9	AL532717	AL532717 AL532717
c 42		152.8	6.1	454	9	AI358948	AI358948 qy24g03.x
43		151.8	6.1	289	10	H33644	H33644 EST109830 R
c 44		151.6	6.0	481	12	AZ025784	AZ025784 RPCI-23-3
45		148	5.9	520	10	BI341423	BI341423 368842 MA

ALIGNMENTS

RESULT 1
BG747345
LOCUS BG747345 819 bp mRNA linear EST 15-MAY-2001
DEFINITION 602704818F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4857941 5', mRNA sequence.
ACCESSION BG747345
VERSION BG747345.1 GI:14057998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 819)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1711 row: d column: 06
High quality sequence stop: 792.
FEATURES Location/Qualifiers
source 1. .819
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4857941"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 202 a 201 c 235 g 181 t
ORIGIN

Query Match 29.5%; Score 738.6; DB 10; Length 819;
Best Local Similarity 96.1%; Pred. No. 3.1e-160;
Matches 789; Conservative 0; Mismatches 29; Indels 3; Gaps 3;

```
Qy    163 GACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGG 222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGG 60

Qy    223 GTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGG 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGG 120

Qy    283 GATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAG 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAG 180

Qy    343 TCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGC 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGC 240

Qy    403 TACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGAT 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGAT 300

Qy    463 TGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATT 522
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    301 TGGAGAGCAGCAGAACCCA-GGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAGATT 359

Qy    523 CGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTG 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 CGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAGTTG 419

Qy    583 CTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACAT 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 CTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACACAT 479

Qy    643 CATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAAC 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 CATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAGAAC 539

Qy    703 ATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCTAAA 762
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    540 ATCACCATGAAGTGGCTGAAGGATAAGCAG-CAATGGATGCCAAGGAGTTCGAACCTAAA 598

Qy    763 GACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCCCCT 822
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    599 GACGTATTGCCCAATGGGGATGGGACCTACCA-GGCTGGATAACCTTGGCTGTACCCCCT 657

Qy    823 GGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATT 882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    658 GGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTCATT 717

Qy    883 GTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATTGCT 942
          ||||||| ||   ||    ||||||||||||      | | |||||||    | | |
Db    718 GTGATCTTGGGAGCCCTCACGTCTGGCACCTAGTCATTGGAAGTCATCCAGTGGAATTGC 777

Qy    943 GTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAAT 983
          |||||||||||||||||| ||||||||||||||||||||
Db    778 TGTTTTGTCGTCATCTTGTTCCTTGGAATTTTGTTCATAAT 818
```

RESULT 2
AK009581
LOCUS AK009581 1723 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310032M04:hemochromatosis, full insert sequence.
ACCESSION AK009581
VERSION AK009581.1 GI:12844462
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library clone:2310032M04.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4 (sites)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 1723)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Gemomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went

through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES Location/Qualifiers

source 1. .1723
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="MGD:MGI:1905246"
/db_xref="taxon:10090"
/clone="2310032M04"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

gene 99. .1178
/gene="Hfe"

CDS 99. .1178
/gene="Hfe"
/note="data source:MGD, source key:MGI:109191, evidence:ISS hemochromatosis putative"
/codon_start=1
/protein_id="BAB26373.1"
/db_xref="GI:12844463"
/translation="MSLSAGLPVRPLLLLLLLLWSVAPQALPPRSHSLRYLFMGASEP DLGLPLFEARGYVDDQLFVSYNHESRRAEPRAPWILEQTSSQLWLHLSQSLKGWDYMF IVDFWTIMGNYNHSKVTKLGVVSESHILQVVLGCEVHEDNSTSGFWRYGYDGQDHLEF CPKTLNWSAAEPGAWATKVEWDEHKIRAKQNRDYLEKDCPEQLKRLLELGRGVLGQQV PTLVKVTRHWASTGTSLRCQALDFFPQNITMRWLKDNQPLDAKDVNPEKVLPNGDETY QGWLTLAVAPGDETRFTCQVEHPGLDQPLTASWEPLQSQAMIIGIISGVTVCAIFLVG ILFLILRKRKASGGTMGGYVLTDCE"

polyA_signal 1695. .1700
/note="putative"

polyA_site 1723
/note="putative"

BASE COUNT 406 a 456 c 454 g 407 t
ORIGIN

Query Match 25.6%; Score 640.8; DB 11; Length 1723;
Best Local Similarity 68.1%; Pred. No. 1.5e-137;
Matches 1121; Conservative 0; Mismatches 437; Indels 88; Gaps 13;

```
Qy      4 GGCCCGCGAGCCAGGCCGGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCAG 63
          || |  |  |   |||||  ||| || || ||| | ||  ||  | ||| || || ||||
Db    114 GGGCTCCCTGTGCGGCCGCTGCTGCTGCTGCTACTGCTGTGGTCCGTGGCCCCGCAG 173

Qy     64 GGGCGCTTGCTGCGTTCACACTCTCTGCACTACCTCTTCATGGGTGCCTCAGAGCAGGAC 123
          |  |     | |||||||| |||||    ||||||||||||||||||||||||  |||
Db    174 GCACTGCCACCGCGTTCACATTCTCTAAGATACCTCTTCATGGGTGCCTCAGAGCCAGAC 233

Qy    124 CTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCGTGTTCTAT 183
          || || ||  | |||||||| |||  |||||| |||||||||||||| || |||| |||
Db    234 CTCGGGCTGCCTTTGTTTGAGGCTAGGGGCTATGTGGATGACCAGCTCTTTGTGTCCTAC 293

Qy    184 GATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGC 243
          ||||||||||||||||||  |||||| |  | || ||| | |        ||  ||||||
Db    294 AATCATGAGAGTCGCCGTGCTGAGCCCAGGGCCCCGTGGATCTTGGAGCAAACCTCAAGC 353

Qy    244 CAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGAC 303
          ||| ||||||||||| ||||||||||| ||||||||||||||  |||||||||  || |||
Db    354 CAGCTGTGGCTGCATCTGAGTCAGAGCCTGAAAGGGTGGGACTACATGTTCATAGTAGAC 413

Qy    304 TTCTGGACTATTATGGAAAATCACAACCACAGCAAG---------------------- 339
          ||||||||  || ||||  ||  | ||||||||  |||
Db    414 TTCTGGACCATCATGGGCAACTATAACCACAGTAAGGTCACGAAGTTGGGAGTGGTGTCC 473

Qy    340 GAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAG 399
          |||||||||| |||||||||  |||| ||||||||  |||| ||||||||||||||||
Db    474 GAGTCCCACATCCTGCAGGTGGTCCTAGGCTGTGAGGTGCATGAAGACAACAGTACCAGC 533

Qy    400 GGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTG 459
          |||| |||||  || || ||||| ||||| || ||||| ||||||||||  | |||||
Db    534 GGCTTCTGGAGATATGGTTATGACGGGCAAGATCACCTGGAATTCTGCCCCAAGACACTA 593

Qy    460 GATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACAAG 519
          | ||||| ||||| || ||  |||||||| |||||||| |||| |||||   |||||||
Db    594 AACTGGAGCGCAGCCGAGCCAGGGGCCTGGGCCACCAAGGTGGAATGGGACGAGCACAAG 653

Qy    520 ATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCAG 579
          || || ||||  |||||||||| |||||||||| |||||||| |  |||||| | | |
Db    654 ATCCGTGCCAAACAGAACAGGGACTACCTGGAGAAGGACTGCCCCGAGCAGCTGAAACGG 713

Qy    580 TTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGACA 639
          | |||||||||||||||| ||| |||  || |||||||| |||||||||| |||||
Db    714 CTCCTGGAGCTGGGGAGAGGCGTTCTGGGACAGCAAGTGCCTACTTTGGTGAAAGTGACT 773

Qy    640 CATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCCAG 699
          |  ||   | ||||| | | |||| |||| |||||| |||  || ||| || |||||||
Db    774 CGCCACTGGGCCTCTACGGGGACCTCTCTAAGGTGTCAGGCTCTGGACTTCTTCCCCCAG 833

Qy    700 AACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAACCT 759
          |||||||| |||| |||| ||||||| || || ||| |||||||||| ||  || | ||
Db    834 AACATCACTATGAGGTGGTTGAAGGACAACCAACCACTGGATGCCAAAGATGTCAACCCC 893

Qy    760 AAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTACCC 819
          |  | ||  | || || ||||||| |||||| || |||||| | || ||||| ||  ||
Db    894 GAGAAGGTGCTACCTAACGGGGATGAGACCTATCAAGGCTGGCTGACATTGGCCGTGGCC 953

Qy    820 CCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCCAGGCCTGGATCAGCCCCTC 879
          |||||||| |||   || |  || || || ||||||||||||||||||||| ||||| |||
Db    954 CCTGGGGACGAGACAAGGTTCACCTGTCAAGTGGAGCACCCAGGCCTGGACCAGCCTCTC 1013

Qy    880 ATTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCATTGGAGTCATCAGTGGAATT 939
          | ||     |||||||||| |  |||    || |  | || ||| |||||||||| |
Db   1014 ACTGCCTCTTGGGAGCCCTTGCAATCTCAGGCCATGATTATCGGAATCATCAGTGGAGTC 1073

Qy    940 GCTGTTTTTGTCGTCATCTTGTTCATTGGAATTTTGTTCATAATATTAAGGAAGAGGCAG 999
          | || | |   | |||||| ||  |||||||| ||||| |||| |||||||| ||| ||
Db   1074 ACCGTCTGT---GCCATCTTCTTGGTTGGAATTCTGTTCCTAATCTTAAGGAAAAGGAAG 1130

Qy   1000 GGTTCAAGAGGAGCCATGGGGCACTACGTCTTAGCTGAACGTGAGTGACACGCAGCCTGC 1059
          | |||| ||||| |||||||   ||| |||||| | ||  ||||||||   |||||||||
Db   1131 GCTTCAGGAGGAACCATGGGTGGCTATGTCTTAACAGACTGTGAGTGATCTGCAGCCTGC 1190

Qy   1060 AGACTCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCT 1119
          ||  ||| |  |  || | ||  | | | |||||   || |||     | |     |
Db   1191 TGAACCACGGAAGAGAGAAAACTCAGCCAAAGACTTGGAGGGGGCACACTTGCTCCA--- 1247

Qy   1120 CTTCATGTTTCAGGAGAGAGTTGAACCTAA-ACATAGAAATTGCCTGACGAACTCCTTGA 1178
                |  |||| | ||||| |||||| ||| ||||| |||||||  |   |  ||
Db   1248 -------CTGTAGGACACAGTTGGACCTAACACACAGAAACTGCCTGAGAA---CTGTGC 1297

Qy   1179 TTTTAGCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCT 1238
          | ||||| |||||||||||| || || ||   | |||| ||| ||| ||   | |||||||
Db   1298 TCTTAGCTTTCTCTGTTCACTTTCTTAAGGTGTTTTCTCCAGTTAAGT---TCAGTTCCT 1354

Qy   1239 GCATGCCGGTGAT--CCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCA 1296
          | ||    |||||  | | || ||  |||||||||      |     ||||||||  || |
Db   1355 GAATAGTAGTGATTGCACCAGTTGCAACCTCTCCCTCCAGAACTGGTCTCATGATTCTTA 1414

Qy   1297 AGCTGCATCTA-GAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTC 1355
          |||||  |||  || || ||||   |||||| | |    |  |||  |
Db   1415 GGCTGCTTCTTGGAAGCATCCTATGTTTCCTTCATGCACCTAGACTC------------- 1461

Qy   1356 ATTTCATTTCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATT 1415
                | ||| | |      | ||  ||  |  |  |    ||||||||||  ||
Db   1462 ---------CATATGTCTACGTAAAGAGCCCCTCTAAGTTTAGTGGATACATGATTCGTT 1512

Qy   1416 TTAACATCTGAGAAAAGCTTTGAACCCTGGGACGTGGCT----AGTCATAACCTTACCAG 1471
          |  |||||||| | ||| | |||||| |     || || |     |  ||||      ||||
Db   1513 TCCACATCTGA-AGAAGTTGTGAACCTTCATCCGGGGATGCTCACACATACTTGAGCCAG 1571
```

```
Qy   1472 ATTTTTACACATGTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTTGAATCCTCT 1531
          | |||| ||| | ||||  | |  | | |||||| ||||||  |              ||
Db   1572 AATTTTTCACCTATATCCTAGAA--TCCAGGACCCACTCAACTAT-----------CCT 1617

Qy   1532 CTCTGTGTTACCCAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCATCTGATTGT 1591
          |  | |||||   ||| |||| ||||||||||| ||| ||     ||| ||||  || |||
Db   1618 CCATCTGTTATAGAGTGACTCCTCTGTCACCATGCCCTGACTTCTCTGCCATTGGAGTGT 1677

Qy   1592 GATGTGAGTTGCACAGCTATGAAGGC 1617
           || |   | |  || | || ||| |
Db   1678 TATATATATGGATCATCAATAAAGCC 1703


RESULT 3
BE272926
LOCUS       BE272926                570 bp    mRNA    linear   EST 13-JUL-2000
DEFINITION  601171213F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:3544803 5',
            mRNA sequence.
ACCESSION   BE272926
VERSION     BE272926.1  GI:9147279
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM240  row: j  column: 04
            High quality sequence stop: 566.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3544803"
                     /clone_lib="NIH_MGC_14"
                     /tissue_type="renal cell adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      140 a    148 c    175 g    107 t
ORIGIN

  Query Match             20.2%;  Score 506.8;  DB 10;  Length 570;
  Best Local Similarity   99.4%;  Pred. No. 1.3e-106;
  Matches  519;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy    339 GGAGTCCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGA 398
          | ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     48 GCAGTCCCACACCGTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGA 107

Qy    399 GGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACT 458
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108 GGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACT 167

Qy    459 GGATTGGAGAGCAGCAGAACCCAG-GGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACA 517
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    168 GGATTGGAGAGCAGCAGAACCCAGTGGCCTGGCCCACCAAGCTGGAGTGGGAAAGGCACA 227

Qy    518 AGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGC 577
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 AGATTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACTGCCCTGCACAGCTGCAGC 287

Qy    578 AGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGA 637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    288 AGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGTGA 347

Qy    638 CACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCC 697
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    348 CACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCGGGCCTTGAACTACTACCCCC 407

Qy    698 AGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAAC 757
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    408 AGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGGATGCCAAGGAGTTCGAAC 467

Qy    758 CTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTAC 817
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    468 CTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTGGATAACCTTGGCTGTAC 527

Qy    818 CCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 859
          |||||||||||||||||||||||||||||||||||||||||||
Db    528 CCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 569


RESULT 4
N93736/c
LOCUS       N93736                  520 bp    mRNA    linear   EST 20-AUG-1996
DEFINITION  zb51g06.s1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
            IMAGE:307162 3' similar to gb:M55053 CYTOCHROME P450 IA2 (HUMAN);,
            mRNA sequence.
ACCESSION   N93736
VERSION     N93736.1  GI:1266045
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 520)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1642   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 328.
FEATURES             Location/Qualifiers
     source          1. .520
                     /organism="Homo sapiens"
                     /db_xref="GDB:1250574"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:307162"
                     /clone_lib="Soares_fetal_lung_NbHL19W"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTTT-3'],
```

double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbHH19W."

BASE COUNT 175 a 94 c 88 g 161 t 2 others
ORIGIN

Query Match 18.0%; Score 451; DB 10; Length 520;
Best Local Similarity 95.4%; Pred. No. 1.1e-93;
Matches 497; Conservative 0; Mismatches 17; Indels 7; Gaps 3;

```
Qy   1985 GTACTATTATCCCCATTTCTTTTTTAAATGAAGAAAGTGAAGTAGGCCGGGCACGG---- 2040
          ||||||||| |||||||||||||||||||| |||||||||||||||||||||||||||
Db    520 GTACTATTANCCCCATTTCTTTTTTAAATGAGGAAAGTGAAGTAGGCCGGGCACGGGTGG 461

Qy   2041 -TGGCTCGCGCCTGTGGTCCCAGGGTGCTGAGATT-GCAGGTGTGAGCCACCCTGCCCAG 2098
           ||  ||  | |   |||||  |||||||||||  |||||||||||||||||||||||||
Db    460 CACGCCTGCCTCGGNCCTCCCAAAGTGCTGAGATTACCAGGTGTGAGCCACCCTGCCCAG 401

Qy   2099 CCGTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACA 2158
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    400 CCGTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACA 341

Qy   2159 TGCACTTGGCTGCATAAATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGA 2218
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    340 TGCACTTGGCTGCATAAATGTGGTACAAGCATTCTGTCTTGAAGGGCAGGTGCTTCAGGA 281

Qy   2219 TACCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCAT 2278
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    280 TACCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCAT 221

Qy   2279 CTCTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAA 2338
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 CTCTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAA 161

Qy   2339 TGTGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAA 2398
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 TGTGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAA 101

Qy   2399 TTAGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTA 2458
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    100 TTAGCCAGTGAAAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTT-CATTA 42

Qy   2459 AAAATGCATATACTTTAATAAATGTACATTGTATTGTAAAA 2499
          |||||||||||||||||||||||| ||||||||||||||
Db     41 AAAATGCATATACTTTAATAAATGTATATTGTATTGTAAAA 1
```

RESULT 5
AI949947/c
LOCUS AI949947 520 bp mRNA linear EST 06-SEP-1999
DEFINITION wq04f05.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2470305 3', mRNA sequence.
ACCESSION AI949947
VERSION AI949947.1 GI:5742257
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 520)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 406.
FEATURES Location/Qualifiers
source 1. .520
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2470305"
/clone_lib="NCI_CGAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 171 a 84 c 82 g 183 t
ORIGIN

Query Match 17.5%; Score 437.8; DB 9; Length 520;
Best Local Similarity 98.4%; Pred. No. 1.2e-90;
Matches 442; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy   2058 CCCAGGGTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCCGTCAAAAGAGTCTTAAT 2117
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     463 CCCAGGGTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCCGTCAAAAGAGTCTTAAT 404

Qy   2118 ATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCATAAAT 2177
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     403 ATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCCGCATAAAT 344

Qy   2178 GTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAGAA 2237
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     343 GTGGTACAAGCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAGAA 284

Qy   2238 GTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCATTT 2297
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     283 GTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCATTT 224

Qy   2298 TCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCATTGT 2357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     223 TCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCATTGT 164

Qy   2358 AGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAAAACTAT 2417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     163 AGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAAAACTAT 104

Qy   2418 TAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATATACTTTAAT 2477
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     103 TAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAACAATGCATATACTTTAAT 44

Qy   2478 AAATGTACATTGTATTGTAAAAAAAAAAA 2506
          |||||||||    ||||||||||||||||
Db      43 AAATGTACAAAAAATTGTAAAAAAAAAAA 15
```

RESULT 6
AW469921/c
LOCUS AW469921 440 bp mRNA linear EST 24-FEB-2000
DEFINITION ha27e08.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2874950 3' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION AW469921
VERSION AW469921.1 GI:7040027

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 439.
FEATURES             Location/Qualifiers
     source          1. .440
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2874950"
                     /clone_lib="NCI_CGAP_Kid12"
                     /tissue_type="2 pooled tumors (clear cell type)"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1323912-1325831, 1471368-1472903 and
                     1492104-1493255). Subtraction by Bento Soares and M.
                     Fatima Bonaldo. "
BASE COUNT      153 a     77 c     68 g    141 t      1 others
ORIGIN

  Query Match             17.4%;  Score 435.8;  DB 9;  Length 440;
  Best Local Similarity   99.3%;  Pred. No. 3.4e-90;
  Matches  437;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   2059 CCAGGGTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCCGTCAAAAGAGTCTTAATA 2118
           ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    440 CCAGGGTGCTGAGATTGCAGGTGTGAGCCACCNTGCCCAGCCGTCAAAAGAGTCTTAATA 381

Qy   2119 TATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCATAAATG 2178
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    380 TATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCCGCATAAATG 321

Qy   2179 TGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAGAAG 2238
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    320 TGGTACAAGCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAGAAG 261

Qy   2239 TTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCATTTT 2298
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    260 TTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCATTTT 201

Qy   2299 CTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCATTGTA 2358
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    200 CTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCATTGTA 141

Qy   2359 GAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAAAACTATT 2418
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    140 GAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAAAACTATT 81

Qy   2419 AACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATATACTTTAATA 2478
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     80 AACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATATACTTTAATA 21

Qy   2479 AATGTACATTGTATTGTAAA 2498
           ||||||||||||||||||||
Db     20 AATGTACATTGTATTGTAAA 1


RESULT 7
AA876054/c
LOCUS       AA876054                439 bp    mRNA    linear   EST 31-MAR-1998
DEFINITION  nr14d03.s1 NCI_CGAP_Co10 Homo sapiens cDNA clone IMAGE:1161797 3'
            similar to contains element MER36 repetitive element ;, mRNA
            sequence.
ACCESSION   AA876054
VERSION     AA876054.1  GI:2984817
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 439)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.
            , Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 865    Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 421.
FEATURES             Location/Qualifiers
     source          1. .439
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1161797"
                     /clone_lib="NCI_CGAP_Co10"
                     /tissue_type="colon tumor RER+"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; 1st strand cDNA was prepared from
                     RER+ colon tumor, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo (N-Soares4). "
BASE COUNT      154 a     77 c     66 g    142 t
ORIGIN

  Query Match             17.3%;  Score 434.2;  DB 9;  Length 439;
  Best Local Similarity   99.3%;  Pred. No. 8e-90;
  Matches  436;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   2064 GTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCCGTCAAAAGAGTCTTAATATATAT 2123
           |||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    439 GTGCTGGGATTGCAGGTGTGGGCCACCCTGCCCAGCCGTCAAAAGAGTCTTAATATATAT 380

Qy   2124 ATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCATAAATGTGGTA 2183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    379 ATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCATAAATGTGGTA 320

Qy   2184 CAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAGAAGTTTCT 2243
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 CAAGCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAGAAGTTTCT 260
```

```
Qy   2244 TCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCATTTTCTTTT 2303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 TCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCATTTTCTTTT 200

Qy   2304 TTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCATTGTAGAAAA 2363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 TTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCATTGTAGAAAA 140

Qy   2364 GCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAAAACTATTAACAA 2423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 GCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAAAACTATTAACAA 80

Qy   2424 CTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATATACTTTAATAAATGT 2483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 CTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATATACTTTAATAAATGT 20

Qy   2484 ACATTGTATTGTAAAAAAA 2502
          |||||||||||||||||||
Db     19 ACATTGTATTGTAAAAAAA 1
```

RESULT 8
AI127651/c
LOCUS AI127651 438 bp mRNA linear EST 27-OCT-1998
DEFINITION qc30h07.x1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone IMAGE:1711165 3' similar to contains element MER36 repetitive element ;, mRNA sequence.
ACCESSION AI127651
VERSION AI127651.1 GI:3596165
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 438)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 669 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 415.
FEATURES Location/Qualifiers
source 1. .438
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1711165"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by M. Fatima Bonaldo."
BASE COUNT 153 a 73 c 67 g 145 t
ORIGIN

Query Match 17.3%; Score 433.8; DB 9; Length 438;
Best Local Similarity 99.5%; Pred. No. 9.9e-90;
Matches 435; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy   2064 GTGCTGAGATTGCAGGTGTGAGCCACCCTGCCCAGCCGTCAAAAGAGTCTTAATATATAT 2123
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    437 GTGCTGAGATTACAGGTGTGAGCCACCCTGCCCAGCCGTCAAAAGAGTCTTAATATATAT 378

Qy   2124 ATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCATAAATGTGGTA 2183
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    377 ATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCAGCATAAATGTGGTA 318

Qy   2184 CAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAGAAGTTTCT 2243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317 CAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAGAAGTTTCT 258

Qy   2244 TCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCATTTTCTTTT 2303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 TCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCATTTTCTTTT 198

Qy   2304 TTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCATTGTAGAAAA 2363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 TTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCATTGTAGAAAA 138

Qy   2364 GCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAAAACTATTAACAA 2423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 GCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAAAACTATTAACAA 78

Qy   2424 CTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATATACTTTAATAAATGT 2483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 CTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATATACTTTAATAAATGT 18

Qy   2484 ACATTGTATTGTAAAAA 2500
          |||||||||||||||||
Db     17 ACATTGTATTGTAAAAA 1
```

RESULT 9
AI122894/c
LOCUS AI122894 392 bp mRNA linear EST 24-SEP-1998
DEFINITION ok31a01.s1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:1509384 3', mRNA sequence.
ACCESSION AI122894
VERSION AI122894.1 GI:3538660
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 392)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 538 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 364.
FEATURES Location/Qualifiers
source 1. .392
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1509384"
/clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares NbHSF pool 1: 309384-310919, 323208-325895 Soares Nb2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 -

150407, 151176-152327 Soares Nb2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NbHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NbHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 147 a 61 c 52 g 132 t
ORIGIN

Query Match 15.5%; Score 388.8; DB 9; Length 392;
Best Local Similarity 99.5%; Pred. No. 2.5e-79;
Matches 390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy   2110 GTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCT 2169
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    392 GTCTTAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCT 333

Qy   2170 GCATAAATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACA 2229
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    332 GCATAAATGTGGTACAAGCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACA 273

Qy   2230 GCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTA 2289
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    272 GCTCAGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTA 213

Qy   2290 AACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACG 2349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    212 AACCATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACG 153

Qy   2350 CTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGA 2409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    152 CTCATTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGA 93

Qy   2410 AAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATAT 2469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     92 AAAACTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATAT 33

Qy   2470 ACTTTAATAAATGTACATTGTATTGTAAAAAA 2501
          |||||||||||||||| ||||||||||||||||
Db     32 ACTTTAATAAATGTATATTGTATTGTAAAAAA 1
```

RESULT 10
BG402460
LOCUS BG402460 805 bp mRNA linear EST 12-MAR-2001
DEFINITION 602466163F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4594359 5',
mRNA sequence.
ACCESSION BG402460
VERSION BG402460.1 GI:13295908
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 805)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1335 row: m column: 16
High quality sequence stop: 606.
FEATURES Location/Qualifiers
source 1. .805
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4594359"
/clone_lib="NIH_MGC_75"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
BASE COUNT 192 a 196 c 178 g 239 t
ORIGIN

Query Match 15.2%; Score 380.2; DB 10; Length 805;
Best Local Similarity 80.3%; Pred. No. 2.8e-77;
Matches 515; Conservative 0; Mismatches 103; Indels 23; Gaps 5;

```
Qy   1223 AGGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTC 1282
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    158 ACGTTTCTGAGTTCCTGCATGCCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTC 217

Qy   1283 TCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGAC 1342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    218 TCTCATGAACCTCAAGCTGCATCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGAC 277

Qy   1343 ATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACT 1402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    278 ATACACCTATGTCATTTCATTTCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACT 337

Qy   1403 TACATGATTCATTTTAACATCTGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAA 1462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    338 TACATGATTCATTTTAACATCTGAGAAAAGCTTTGAACCCTGGGACGTGGCTAGTCATAA 397

Qy   1463 CCTTACCAGATTTTTACACATGTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTT 1522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    398 CCTTACCAGATTTTTACACATGTATCTATGCATTTTCTGGACCCGTTCAACTTTTCCTTT 457

Qy   1523 GAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCACCAAGCCTTGGGGATTCTTCCA 1582
          |||||||||||||||||||||||||||||||||||||||||||||||| ||| | | | |
Db    458 GAATCCTCTCTCTGTGTTACCCAGTAACTCATCTGTCACCAAGCCTGGGGATTCTTCCAA 517

Qy   1583 TCTGATTGTGATGTGAGTTG-----CACAGCTATGAAGGCTGTGCACTGCACGAATGGAA 1637
          ||||||||||||    |        |||||||| |     || | | | | | | |
Db    518 TCTGATTGTGATTGTCGAGTTTGCACACAGCTATTGAAAGGCTGTGCCACTGGCACGAAT 577

Qy   1638 GAGGCACCTGTCCCAG----------AAAAAGCATCATGGCTATCTGTGGGTAGTATGAT 1687
          ||| |  |  || |           |||     ||||||| ||||||||||||||||||
Db    578 GAGAAAGAGGCACCTGGTCCCCAGAAAAAGCCATTCATGGCAATCTGTGGGTAGTATGAT 637

Qy   1688 GGGTGTTTTTAGCA---GGTAGGAGGCAAATATCTTGAAA----GGGGTTGTGAAGAGGT 1740
          ||   |||   |    | |||||||| ||||      |    || ||   ||  ||||
Db    638 CGGCTGTTTATACCCCGGATAGGAGGCACATATTCCTTGAACAGGGCGTGCGGACCAGGT 697

Qy   1741 GTTTTTTCTAATTGGCATGAAGGTGTCATACAGATTTGCAAAGTTTAATGGTGCCTTCAT 1800
          | | |||||   |||||||| ||| ||  |   || ||||  |  | ||| |||| |||
Db    698 GGTATTTCTCTCTGGCATGACGGTTTCTACCGACTTCGCAAGTTACACTGGGGCCTACAT 757

Qy   1801 TTGGGATGCTACTCTAGTATTCCAGACCTGAAGAATCACAA 1841
          | || |  | || |||  |||||| |||  |  |  |||||
Db    758 TGGGCAGACAAC-CTAAGATTCCACACCGAAGAATACACAA 797
```

RESULT 11
AI040303/c
LOCUS AI040303 380 bp mRNA linear EST 24-SEP-1998
DEFINITION oy09c09.x1 Soares_senescent_fibroblasts_NbHSF Homo sapiens cDNA
clone IMAGE:1665328 3', mRNA sequence.
ACCESSION AI040303
VERSION AI040303.1 GI:3279497
KEYWORDS EST.
SOURCE human.

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 380)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 874   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .380
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1665328"
                     /clone_lib="Soares_senescent_fibroblasts_NbHSF"
                     /tissue_type="senescent fibroblast"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker V_TYPE: phagemid; Site_1: Not I; Site_2: Eco RI
                     ; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      143 a     60 c     54 g    123 t
ORIGIN

  Query Match             15.0%;  Score 375.2;  DB 9;  Length 380;
  Best Local Similarity   99.2%;  Pred. No. 3.6e-76;
  Matches  377;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy    2116 ATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCATAA 2175
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     380 ATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCATAA 321

Qy    2176 ATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAG 2235
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     320 ATGTGGTACAAGCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTCAG 261

Qy    2236 AAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCAT 2295
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 AAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACCAT 201

Qy    2296 TTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCATT 2355
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 TTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCATT 141

Qy    2356 GTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAAAACT 2415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 GTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAAAACT 81

Qy    2416 ATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATATACTTTA 2475
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      80 ATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATATACTTTA 21

Qy    2476 ATAAATGTACATTGTATTGT 2495
           ||  ||||||||||||||||
Db      20 ATCCATGTACATTGTATTGT 1


RESULT 12
AI763178/c
LOCUS       AI763178                393 bp    mRNA    linear   EST 24-JUN-1999
DEFINITION  wi65f02.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2398203 3',
            mRNA sequence.
ACCESSION   AI763178
VERSION     AI763178.1  GI:5178845
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 393)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .393
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2398203"
                     /clone_lib="NCI_CGAP_Kid12"
                     /tissue_type="2 pooled tumors (clear cell type)"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1323912-1325831, 1471368-1472903 and
                     1492104-1493255). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
BASE COUNT      142 a     62 c     53 g    136 t
ORIGIN

  Query Match             14.6%;  Score 365.8;  DB 9;  Length 393;
  Best Local Similarity   95.7%;  Pred. No. 5.3e-74;
  Matches  376;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy    2114 TAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCAT 2173
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     393 TAATATATATATCCAGATGGCATGTGTTTACTTTATGTTACTACATGCACTTGGCTGCAT 334

Qy    2174 AAATGTGGTACAACCATTCTGTCTTGAAGGGCAGGTGCTTCAGGATACCATATACAGCTC 2233
           |||||||||||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db     333 AAATGTGGTACAAGCATTCTGTCTTGAAGGGCAGGTGCTTCAGAATACCATATACAGCTC 274

Qy    2234 AGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAAACC 2293
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     273 AGAAGTTTCTTCTTTAGGCATTAAATTTTAGCAAAGATATCTCATCTCTTCTTTTAACCC 214

Qy    2294 ATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCA 2353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     213 ATTTTCTTTTTTTGTGGTTAGAAAAGTTATGTAGAAAAAAGTAAATGTGATTTACGCTCA 154

Qy    2354 TTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAAAA 2413
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     153 TTGTAGAAAAGCTATAAAATGAATACAATTAAAGCTGTTATTTAATTAGCCAGTGAAGAA 94

Qy    2414 CTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAAAAATGCATATACTT 2473
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |||
Db      93 CTATTAACAACTTGTCTATTACCTGTTAGTATTATTGTTGCATTAACAATGCATATGCTT 34
```

```
Qy    2474 TAATAAATGTACATTGTATTGTAAAAAAAAAAA 2506
           ||||||||||          |   ||||||||||||
Db      33 TAATAAATGTTAGAGAATTGTAAAAAAAAAAAA 1


RESULT 13
BF883952/c
LOCUS       BF883952                384 bp    mRNA    linear   EST 17-JAN-2001
DEFINITION  PM4-ET0209-151200-003-f07 ET0209 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF883952
VERSION     BF883952.1  GI:12274078
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 384)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM4&t2=PM4-ET0209-
            151200-003-f07&t3=2000-12-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 17
            High quality sequence stop: 384.
FEATURES             Location/Qualifiers
     source          1. .384
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="ET0209"
                     /dev_stage="Adult"
                     /note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       92 a    112 c     87 g     93 t
ORIGIN

  Query Match           14.3%;  Score 358;  DB 10;  Length 384;
  Best Local Similarity   99.7%;  Pred. No. 3.4e-72;
  Matches  369;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy     104 TGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATG 163
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     384 TGGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTTGTTTGAAGCTTTGGGCTACGTGGATG 325

Qy     164 ACCAGCTGTTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGG 223
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     324 ACCAGCTG-TCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCCGAACTCCATGGG 266

Qy     224 TTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGG 283
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     265 TTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAAAGGGTGGG 206

Qy     284 ATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGT 343
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 ATCACATGTTCACTGTTGACTTCTGGACTATTATGGAAAATCACAACCACAGCAAGGAGT 146

Qy     344 CCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCT 403
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145 CCCACACCCTGCAGGTCATCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCT 86

Qy     404 ACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATT 463
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85 ACTGGAAGTACGGGTATGATGGGCAGGACCACCTTGAATTCTGCCCTGACACACTGGATT 26

Qy     464 GGAGAGCAGC 473
           ||||||||||
Db      25 GGAGAGCAGC 16


RESULT 14
BF739151
LOCUS       BF739151                368 bp    mRNA    linear   EST 10-JAN-2001
DEFINITION  CM4-KT0035-261200-609-b04 KT0035 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF739151
VERSION     BF739151.1  GI:12065827
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 368)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM4&t2=CM4-KT0035-
            261200-609-b04&t3=2000-12-26&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 29
            High quality sequence stop: 366.
FEATURES             Location/Qualifiers
     source          1. .368
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="KT0035"
                     /dev_stage="Adult"
                     /note="Organ: bladder_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       96 a     89 c     71 g    112 t
ORIGIN

  Query Match           13.6%;  Score 341.2;  DB 10;  Length 368;
  Best Local Similarity   98.6%;  Pred. No. 2.6e-68;
```

```
  Matches  365;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy   1064 TCACTGTGGGAAGGAGACAAAACTAGAGACTCAAAGAGGGAGTGCATTTATGAGCTCTTC 1123
          ||||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db      1 TCACTGTGGGAAGGAGACAAAA-TAGAGACTCAAAGAGGGAGTGCATTTATGCGCTCTTC 59

Qy   1124 ATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAATTGCCTGACGAACTCCTTGATTTTA 1183
          |||| |||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db     60 ATGTATCAGGAGAGAGTTGAACCTAAACATAGAAA-TGCCTGACGAACTCCTTGATTTTA 118

Qy   1184 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 1243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    119 GCCTTCTCTGTTCATTTCCTCAAAAAGATTTCCCCATTTAGGTTTCTGAGTTCCTGCATG 178

Qy   1244 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    179 CCGGTGATCCCTAGCTGTGACCTCTCCCCTGGAACTGTCTCTCATGAACCTCAAGCTGCA 238

Qy   1304 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 1363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    239 TCTAGAGGCTTCCTTCATTTCCTCCGTCACCTCAGAGACATACACCTATGTCATTTCATT 298

Qy   1364 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTTACATGATTCATTTTAACATC 1423
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    299 TCCTATTTTTGGAAGAGGACTCCTTAAATTTGGGGGACTCACATGATTCATTTTAACATC 358

Qy   1424 TGAGAAAAGC 1433
          ||||||||||
Db    359 TGAGAAAAGC 368


RESULT 15
BF445847/c
LOCUS       BF445847                388 bp    mRNA    linear   EST 01-DEC-2000
DEFINITION  7p38b08.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3647966 3',
            mRNA sequence.
ACCESSION   BF445847
VERSION     BF445847.1  GI:11510985
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 388)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 380.
FEATURES             Location/Qualifiers
     source          1. .388
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3647966"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      118 a     77 c     73 g    109 t     11 others
ORIGIN

  Query Match           13.3%;  Score 333.8;  DB 10;  Length 388;
  Best Local Similarity   93.8%;  Pred. No. 1.4e-66;
  Matches  364;  Conservative    0;  Mismatches   18;  Indels    6;  Gaps    2;

Qy   2029 GGCCGGGCACGGTGGCTCGCGCCTGTGG-----TCCCAGGGTGCTGAGATTGCAGGTGTG 2083
          |||||||| ||| |  || |||| |         |||||||| | | | |  ||||||||
Db    388 GGCCGGGCNCGGGGCTTCACGCCNGCCNCGGCCTCCCAGGGNGNTNANANNGCAGGTGTN 329

Qy   2084 AGCCA-CCCTGCCCAGCCGTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTT 2142
          | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 ANCCACCCCTGCCCAGCCGTCAAAAGAGTCTTAATATATATATCCAGATGGCATGTGTTT 269

Qy   2143 ACTTTATGTTACTACATGCACTTGGCTGCATAAATGTGGTACAACCATTCTGTCTTGAAG 2202
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    268 ACTTTATGTTACTACATGCACTTGGCTGCATAAATGTGGTACAAGCATTCTGTCTTGAAG 209

Qy   2203 GGCAGGTGCTTCAGGATACCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTT 2262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 GGCAGGTGCTTCAGGATACCATATACAGCTCAGAAGTTTCTTCTTTAGGCATTAAATTTT 149

Qy   2263 AGCAAAGATATCTCATCTCTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTA 2322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148 AGCAAAGATATCTCATCTCTTCTTTTAAACCATTTTCTTTTTTTGTGGTTAGAAAAGTTA 89

Qy   2323 TGTAGAAAAAAGTAAATGTGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAAT 2382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88 TGTAGAAAAAAGTAAATGTGATTTACGCTCATTGTAGAAAAGCTATAAAATGAATACAAT 29

Qy   2383 TAAAGCTGTTATTTAATTAGCCAGTGAA 2410
          ||||||||||||||||||||||||||||
Db     28 TAAAGCTGTTATTTAATTAGCCAGTGAA 1


Search completed: November  2, 2002, 06:42:06
Job time : 3804.46 secs
```